

SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: \_\_\_\_\_ Examiner #: \_\_\_\_\_ Date: \_\_\_\_\_  
Art Unit: \_\_\_\_\_ Phone Number 30 \_\_\_\_\_ Serial Number: \_\_\_\_\_  
Mail Box and Bldg/Room Location: \_\_\_\_\_ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need:

\*\*\*\*\*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: \_\_\_\_\_

Inventors (please provide full names): \_\_\_\_\_

Earliest Priority Filing Date: \_\_\_\_\_

*\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

FOR OFFICIAL USE ONLY

STAFF USE ONLY

Searcher: B. Smith  
Searcher Phone #: 308-4477  
Searcher Location: CM1-1E17  
Date Searcher Picked Up: 2/15/01  
Date Completed: 3/1/01  
Searcher Prep & Review Time: \_\_\_\_\_  
Clerical Prep Time: 3  
Online Time: \_\_\_\_\_

Type of Search

NA Sequence (#) 2  
AA Sequence (#) 2  
Structure (#) \_\_\_\_\_  
Bibliographic \_\_\_\_\_  
Litigation \_\_\_\_\_  
Fulltext \_\_\_\_\_  
Patent Family \_\_\_\_\_  
Other \_\_\_\_\_

Vendors and cost where applicable

STN \_\_\_\_\_  
Dialog \_\_\_\_\_  
Questel/Orbit \_\_\_\_\_  
Dr.Link \_\_\_\_\_  
Lexis/Nexis \_\_\_\_\_  
Sequence Systems 05502  
WWW/Internet \_\_\_\_\_  
Other (specify) \_\_\_\_\_

**THIS PAGE BLANK (USPIC)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2001, 19:53:28 ; Search time 119.29 Seconds  
(without alignments)  
2220.154 Million cell updates/sec

Title: US-09-320-713-1

Perfect score: 705

Sequence: 1 ggcacgagcggacacgcatg.....aaaaaaaaaaaaaaaaaaaaa 705

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq\_36:\*

- 1: /SIDS6/gcgdata/geneseq/geneseqn/NA1980.DAT:\*
- 2: /SIDS6/gcgdata/geneseq/geneseqn/NA1981.DAT:\*
- 3: /SIDS6/gcgdata/geneseq/geneseqn/NA1982.DAT:\*
- 4: /SIDS6/gcgdata/geneseq/geneseqn/NA1983.DAT:\*
- 5: /SIDS6/gcgdata/geneseq/geneseqn/NA1984.DAT:\*
- 6: /SIDS6/gcgdata/geneseq/geneseqn/NA1985.DAT:\*
- 7: /SIDS6/gcgdata/geneseq/geneseqn/NA1986.DAT:\*
- 8: /SIDS6/gcgdata/geneseq/geneseqn/NA1987.DAT:\*
- 9: /SIDS6/gcgdata/geneseq/geneseqn/NA1988.DAT:\*
- 10: /SIDS6/gcgdata/geneseq/geneseqn/NA1989.DAT:\*
- 11: /SIDS6/gcgdata/geneseq/geneseqn/NA1990.DAT:\*
- 12: /SIDS6/gcgdata/geneseq/geneseqn/NA1991.DAT:\*
- 13: /SIDS6/gcgdata/geneseq/geneseqn/NA1992.DAT:\*
- 14: /SIDS6/gcgdata/geneseq/geneseqn/NA1993.DAT:\*
- 15: /SIDS6/gcgdata/geneseq/geneseqn/NA1994.DAT:\*
- 16: /SIDS6/gcgdata/geneseq/geneseqn/NA1995.DAT:\*
- 17: /SIDS6/gcgdata/geneseq/geneseqn/NA1996.DAT:\*
- 18: /SIDS6/gcgdata/geneseq/geneseqn/NA1997.DAT:\*
- 19: /SIDS6/gcgdata/geneseq/geneseqn/NA1998.DAT:\*
- 20: /SIDS6/gcgdata/geneseq/geneseqn/NA1999.DAT:\*
- 21: /SIDS6/gcgdata/geneseq/geneseqn/NA2000.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	704.2	99.9	705	21	Partial nucleotide
2	698.6	99.1	1067	21	Nucleotide sequenc
3	640.4	90.8	1047	21	Human Interleukin
4	256.4	36.4	591	21	Human Interleukin-
5	230	32.6	230	21	Virtual DNA fragme
6	141.2	20.0	332	21	CDNA clone HTGEDI9
7	67	9.5	1453	14	Asparaginylendopep
8	67	9.5	1640	14	Asparaginylendopep
9	67	9.5	1910	14	Asparaginylendopep
10	65.8	9.3	2010	21	CDNA encoding huma
11	61.4	8.7	1554	20	Xenopus WA545 prot
12	61.2	8.7	882	18	Alzheimer's diseas

13	61.2	8.7	4237	19	V61487	Human secreted pro
14	61	8.7	1181	19	V59803	Human secreted pro
15	61	8.7	1212	19	V59686	Human secreted pro
16	60.6	8.6	1135	20	X80676	Clone yb7_1 encodi
17	59.2	8.4	281	21	A00286	Human colon cancer
18	59	8.4	1527	21	A16633	Human secreted pro
19	58.8	8.3	1037	21	A26447	Human secreted pro
20	58.6	8.3	1740	21	Z36914	DNA encoding an ac
21	58.6	8.3	1801	21	Z36983	CDNA encoding an a
22	58.4	8.3	2665	19	V32592	Schwannoma1-bindin
23	58.2	8.3	491	21	Z52570	Human secreted pro
24	58.2	8.3	1602	18	T43204	Human mitogen-acti
25	58.2	8.3	1602	20	X07066	Human mitogen acti
26	58	8.2	550	19	V41916	Nucleotide sequenc
27	58	8.2	1212	20	Z28298	Rat neuronal limed
28	58	8.2	1231	20	Z28285	Rat neuronal limed
29	57.8	8.2	1963	21	Z98097	Human secreted pro
30	57.6	8.2	1558	17	T28255	Protein PRO244 CDN
31	57.6	8.2	1560	17	T18828	Survival motor neu
32	57.6	8.2	1582	17	T28259	Survival motor neu
33	57.6	8.2	1582	17	T18831	Human survival mot
34	57.6	8.2	1582	17	V40524	Human survival mot
35	57.6	8.2	2214	19	V64423	Homo sapiens CO122
36	57.4	8.1	467	20	V64423	Mouse developing 1
37	57.4	8.1	916	18	T65651	CDNA encoding anti
38	57.4	8.1	916	18	V62000	R. prolixus NO-r p
39	57.2	8.1	1410	12	Q15022	Hyoscyamine 6 beta
40	57.2	8.1	6200	21	Z49493	Human alpha-L-idur
41	57	8.1	615	18	V02881	Human HMG1-C aberr
42	57	8.1	1701	21	Z64982	Membrane-bound pro
43	56.8	8.1	1965	21	Z98033	Human secreted pro
44	56.8	8.1	2136	21	Z65065	Membrane-bound pro
45	56.6	8.0	2300	20	Z00686	Human GPC3 DNA. H

#### ALIGNMENTS

RESULT 1	
236834	236834 standard; cDNA; 705 BP.
AC	236834;
XX	
DT	13-MAR-2000 (first entry)
DE	Partial nucleotide sequence of human interleukin-21.
XX	
KW	Human; interleukin-22; IL-22; IL-21; immune system disorder;
KW	immune cell chemotaxis; haematopoietic cell disorder;
KW	haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;
KW	respiratory problem; organ rejection; graft-versus-host disease; GVHD;
KW	inflammation; hyperproliferative disorder; tissue regeneration;
KW	embryonic stem cell differentiation; embryonic stem cell proliferation;
KW	haematopoietic lineage; allergic asthma; ss.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	2..265
FT	CDS
FT	Location/Qualifiers
FT	/*tag= a
FT	/product= "partial IL-21"
XX	
PN	W09961617-A1.
XX	
PD	02-DEC-1999.
XX	
PF	27-MAY-1999;
XX	99WO-US11644.
XX	
PR	29-MAY-1998;
PR	98US-0087340.
PR	10-SEP-1998;
PR	98US-0099805.
XX	30-APR-1999;
XX	99US-0131965.





CC of haematopoietic cells, to modulate haemostatic or thrombolytic  
 CC activity, in treating or detecting autoimmune disorders, treating  
 CC asthma (particularly allergic asthma) or other respiratory problems,  
 CC to treat and/or prevent organ rejection or graft-versus-host disease  
 CC (GVHD), to modulate inflammation, to treat or detect hyperproliferative  
 CC disorders, to treat or detect infectious agents, to differentiate,  
 CC proliferate and attract cells, leading to the regeneration of tissues,  
 CC IL-21 and IL-22 may also increase or decrease the differentiation or  
 CC proliferation of embryonic stem cells and haematopoietic lineage, may  
 CC be used to modulate mammalian characteristics.

XX  
 SQ Sequence 1067 BP; 215 A; 371 C; 273 G; 206 T; 2 other;

Query Match 99.1%; Score 698.6; DB 21; Length 1067;  
 Best Local Similarity 99.9%; Pred. No. 1e-124;  
 Matches 700; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 5 cgagtgacacgagatgagacccgctatccacagaagctgcttcgacgagtcgtgc 64
   |||||||
DB 367 cgggtagacacgagatgagacccgctatccacagaagctgcttcgacgagtcgtgc 426
OY 65 agagcgtgctatcgatgcacgagcggcgccgagacagctgcgtcaactcgtgcgctg 124
   |||||||
DB 427 agagcgtgctatcgatgcacgagcggcgccgagacagctgcgtcaactcgtgcgctg 486
OY 125 ctccagagcctgctggtgtgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 184
   |||||||
DB 487 ctccagagcctgctggtgtgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 546
OY 185 aacactgggaccttgacctccacacacgagctcatccacagctcccgctgcgcacatgc 244
   |||||||
DB 547 aacactgggaccttgacctccacacacgagctcatccacagctcccgctgcgcacatgc 606
OY 245 gtgctgcccgtctcagtgtagccgcgaagccgtggggcctctagactggaacagctg 304
   |||||||
DB 607 gtgctgcccgtctcagtgtagccgcgaagccgtggggcctctagactggaacagctg 666
OY 305 tcccagagagggaccccccattatgtatattatgtattatattatattatattatatt 364
   |||||||
DB 667 tcccagagagggaccccccattatgtatattatgtattatattatattatattatatt 726
OY 365 ctacccttggtgtcgtgacatcccgctgtcgtgagagacagcccccactgtctcctca 424
   |||||||
DB 727 ctacccttggtgtcgtgacatcccgctgtcgtgagagacagcccccactgtctcctca 786
OY 425 tctcagcctcagtagtgggggtwgaagagctcagcacctcttccagcccttaagct 484
   |||||||
DB 787 tctcagcctcagtagtgggggtwgaagagctcagcacctcttccagcccttaagct 846
OY 485 ggaagaaagtggtcacaacgctgctgtaccttggtccctgtcctgtcctccggctccc 544
   |||||||
DB 847 ggaagaaagtggtcacaacgctgctgtaccttggtccctgtcctgtcctccggctccc 906
OY 545 ttaccctatcatctgagcctcagggcccccagagctgctcttcccaacccctcttggaagta 604
   |||||||
DB 907 ttaccctatcatctgagcctcagggcccccagagctgctcttcccaacccctcttggaagta 966
OY 605 cccctgttcttaacaattatttaagtgtagctgtattatatttaacatgtagaacacaaa 664
   |||||||
DB 967 cccctgttcttaacaattatttaagtgtagctgtattatatttaacatgtagaacacaaa 1026
OY 665 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
   |||||||
DB 1027 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1067

```

RESULT 3  
 ID 229728  
 XX 229728 standard; cDNA; 1047 BP.  
 AC 229728;  
 XX

DT 27-MAR-2000 (first entry)  
 DE Human Interleukin 17C, PRO1122 cDNA.  
 XX  
 KW Interleukin; IL-17C; PRO1122 polypeptide; clone DNA62377-1381-1; UNQ561;  
 KW cytokine IL-17; cytotoxic T-lymphocyte-associated antigen 8; CTRA-8;  
 KW hybridisation probe; antagonist; degenerative cartilaginous disorder;  
 KW agonist; diagnosis; therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 Key Location/Qualifiers  
 FH sig\_peptide 50..103  
 FT mat\_peptide 50..634  
 FT /\*tag= a  
 FT /\*tag= b  
 FT /label= Mature\_Interleukin\_17C\_polypeptide  
 FT /note= "Designated as clone DNA62377-1381-1"

WO9960127-A2.  
 XX  
 XX 25-NOV-1999.  
 XX  
 PD 14-MAY-1999; 99WO-US10733.  
 XX  
 PE 15-MAY-1998; 98US-0085579.  
 PR 23-DEC-1998; 98US-0113621.  
 XX  
 PR (GENTH ) GENENTECH INC.  
 PA  
 PI Chen J, Filvaroff E, Goddard A, Gurney AL, Li H, Wood WL;  
 XX  
 DR WPT. 2000-116314/10.  
 DR P-PSDB; Y44460.

PT New polypeptides designated PRO1031 and PRO1122 used to treat a  
 PT degenerative cartilaginous disorder -  
 XX  
 PS Claim 2; Fig 4; 141pp; English.

XX  
 XX The present sequence is the cDNA clone DNA62377-1381-1, encoding the  
 CC human PRO1122 polypeptide, also referred to as UNQ561, and as  
 CC interleukin-17C (IL-17C). This sequence has identity with the  
 CC cytokine IL-17 and cytotoxic T-lymphocyte-associated antigen 8 (CTLA-8).  
 CC PRO1122 is expressed in pancreas, small intestine, stomach and testis  
 CC also. It shares about 26-28% amino acid identity with IL-17 and IL-17B.  
 CC The entire coding region of IL-17C can be used as hybridisation probe.  
 CC The PRO1122 polypeptide, agonist or antagonist, is used to diagnose and  
 CC treat a degenerative cartilaginous disorder.

XX  
 SQ Sequence 1047 BP; 177 A; 381 C; 280 G; 209 T; 0 other;

Query Match 90.8%; Score 640.4; DB 21; Length 1047;  
 Best Local Similarity 98.3%; Pred. No. 1.2e-113;  
 Matches 655; Conservative 2; Mismatches 8; Indels 1; Gaps 1;

```

OY 5 cgagtgacacgagatgagacccgctatccacagaagctgcttcgacgagtcgtgc 64
   |||||||
DB 383 cgtgtagacacgagatgagacccgctatccacagaagctgcttcgacgagtcgtgc 442
OY 65 agagcgtgctatcgatgcacgagcggcgccgagacagctgcgtcaactcgtgcgctg 124
   |||||||
DB 443 agagcgtgctatcgatgcacgagcggcgccgagacagctgcgtcaactcgtgcgctg 502
OY 125 ctccagagcctgctggtgtgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 184
   |||||||
DB 503 ctccagagcctgctggtgtgtgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 562
OY 185 aacactgggaccttgacctccacacacgagctcatccacagctcccgctgcgcacatgc 244
   |||||||
DB 563 aacactgggaccttgacctccacacacgagctcatccacagctcccgctgcgcacatgc 622

```

```
OY 245 gtgtgtcccgcttcagtgtgacgcgccaaggccgtggtggcccttagactgtgacacgtgtgc 304
    |||||||
Db 623 gtgtgtcccgcttcagtgtgacgcgccaaggccgtggtggcccttagactgtgacacgtgtgc 682
OY 305 tccccagagggcaccgccctattatgtgtattatgtttattatatagtccctcccaaca 364
    |||||||
Db 683 tccccagagggcaccgccctattatgtgtattatgtttattatatagtccctcccaaca 742
OY 365 ctaccctgtgtgtggtgcatctcccgctgtctgtagagacaagcccccacgtgtctctca 424
    |||||||
Db 743 ctaccctgtgtgtggtgcatctcccgctgtctgtagagacaagcccccacgtgtctctca 802
OY 425 tctccagctcagtagtgtgggtgtwgaagagagctcagacaccttcagcccttaagct 484
    |||||||
Db 803 tctccagctcagtagtgtgggtgtwgaagagagctcagacaccttcagcccttaagct 862
OY 485 gcacaaaggtgtacacacggtgtgctgtacctgtgttccctgtctctgtcccgcttccc 544
    |||||||
Db 863 gcagaaaggtgtacacacggtgtgctgtacctgtgttccctgtctctgtcccgcttccc 922
OY 545 ttaccctatcagctgacctcagagcccgcaagctgacctcttcccaacctcttggaagta 604
    |||||||
Db 923 ttaccctatcagctgacctcagagcccgcaagctgacctcttcccaacctcttggaagta 981
OY 605 cccctgttcttaacaattatttaagtgtacgtgtattataactgtatgacacacaaa 664
    |||||||
Db 982 cccctgttcttaacaattatttaagtgtacgtgtattataactgtatgacacacacac 1041
OY 665 aaaaaa 670
    |||||
Db 1042 ccaaaa 1047

RESULT 4
ID A09153 standard; DNA; 591 BP.
AC A09153;
XX
XX 10-AUG-2000 (first entry)
DT
DE Human interleukin-17 (IL-17) homologue coding sequence.
XX
KW Interleukin 17; IL-17; haematopoiesis; chemotherapy; cytostatic;
KW antianemic; cardiast; hemostatic; anti-inflammatory; anti-HIV; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT sig_peptide 1..54
FT /*tag= a
FT mat_peptide 55..591
FT /*tag= b
XX
XX W0200020593-A1.
XX
XX 13-APR-2000.
XX
XX 30-SEP-1999; 99WO-US22678.
XX
XX 02-OCT-1998; 98US-0102883.
XX 01-DEC-1998; 98US-0110405.
XX 11-JUN-1999; 99US-0138910.
XX
XX (ELIL ) LILLY & CO ELI.
XX
XX Glasebrook AL, Su EW, We1 J, Liu L;
XX
XX WPI: 2000-303778/26.
XX
XX P-PSDB: Y92238.
XX
XX Nucleic acid encoding an interleukin-17 (IL-17) homolog polypeptide
XX which enhances hematopoiesis, useful for treating e.g. anemia,
```

```
PT thrombocytopenia, viral and bacterial infections
XX
XX Claim 6; Page 92; 111pp; English.
XX
CC Interleukin 17 (IL-17) stimulates hematopoiesis and production of
CC neutrophils, granulocytes, or platelets, this may be useful during
CC chemotherapy. IL-17 homologues have at least one actively selected
CC from induction of cytotoxic T cells, induction of lymphokine-activated
CC killer cell proliferation or a B or T cell stimulation. The IL-17
CC homologue may also be used to treat viral or bacterial infections.
CC immune related diseases, anemia, leukemia, thrombocytopenia, uremia,
CC Von Willebrand disease, postoperative cardiovascular dysfunction,
CC treatment of AIDS (acquired immune deficiency syndrome)-related bone
CC marrow failure, and inflammatory diseases of the gastrointestinal
CC system, joints, and lungs.
XX
SQ Sequence 591 BP; 91 A; 223 C; 179 G; 98 T; 0 other;

Query Match 36.4%; Score 256.4; DB 21; Length 591;
Best Local Similarity 99.6%; Pred. No. 1,1e-40;
Matches 257; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY . 5 cgagtgcacacgagatgagacgcgtatccacaagaagctggccttcgcgagtgctgtgc 64
    |||||||
Db 334 cgtgtgacacagatgagacgcgtatccacaagaagctggccttcgcgagtgctgtgc 393
OY 65 agagcgtgtatcgatgcagcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 124
    |||||||
Db 394 agagcgtgtatcgatgcagcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 453
OY 125 ctccagagcgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 184
    |||||||
Db 454 ctccagagcgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 513
OY 185 acacctgggaccttgccttcacacacgagttcatccagttcccgctggctgacctgc 244
    |||||||
Db 514 acacctgggaccttgccttcacacacgagttcatccagttcccgctggctgacctgc 573
OY 245 gtgtgtcccgcttcagtg 262
    |||||||
Db 574 gtgtgtcccgcttcagtg 591

RESULT 5
ID 229731 standard; DNA; 230 BP.
XX
XX 229731;
XX
XX 27-MAR-2000 (first entry)
DT
DE Virtual DNA fragment DNA49665, for isolation of human PRO1122 cDNA.
XX
KW Interleukin; IL-17C; PRO1122 polypeptide; clone DNA62377-1381-1; UN0561;
KW PCR primer; probe; antagonist; degenerative cartilaginous disorder;
KW agonist; virtual DNA fragment; DNA49665; ss.
XX
XX Synthetic.
XX
XX W09960127-A2.
XX
XX 25-NOV-1999.
XX
XX 14-MAY-1999; 99WO-US10733.
XX
XX 15-MAY-1998; 98US-0085579.
XX 23-DEC-1998; 98US-0113621.
XX
XX (GETH ) GENENTECH INC.
XX
XX Chen J, Flivaroff E, Goddard A, Gurney AL, Li H, Wood WI;
```

DR WPI: 2000-116314/10.  
 XX  
 PT New polypeptides designated PRO1031 and PRO1122 used to treat a  
 XX degenerative cartilaginous disorder -  
 XX  
 PS Example 2; Fig 6; 141pp; English.  
 XX  
 CC The present sequence is the virtual DNA fragment DNA4965, used for the  
 CC isolation of the clone DNA62377-1381-1, encoding human PRO1122  
 CC polypeptide, IL-17C. This sequence is used to generate oligonucleotides  
 CC for use as PCR primers to identify the cDNA library that contains the  
 CC sequence of interest and for use as probes for isolation of clone with  
 CC the full-length coding sequence of PRO1122. The PRO1122 polypeptide,  
 CC agonist or antagonist, can be used to diagnose and treat a degenerative  
 CC cartilaginous disorder.  
 CC  
 XX Sequence 230 BP; 32 A; 88 C; 70 G; 40 T; 0 other;  
 SQ  
 Query Match 32.6%; Score 230; DB 21; Length 230;  
 Best Local Similarity 100.0%; Pred. No. 1e-35;  
 Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 13 caccgattgagagccgctatccacagaagctgcttcgagctgctgtgacagagctg 72  
 Db 1 caccgattgagagccgctatccacagaagctgcttcgagctgctgtgacagagctg 60  
 QY 73 tatcagatgacagagcggcgccgagagagctgcttcaactcgtgctgctgtccagag 132  
 Db 61 tatcagatgacagagcggcgccgagagagctgcttcaactcgtgctgctgtccagag 120  
 QY 133 cctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 192  
 Db 121 cctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 180  
 QY 193 ggccttgccttcacacagcagcttccacagctcccgctgctgctgctgctgctgct 242  
 Db 181 ggccttgccttcacacagcagcttccacagctcccgctgctgctgctgctgctgct 230  
 RESULT 6  
 Z36856 236856 standard; DNA; 332 BP.  
 AC Z36856;  
 XX  
 DT 13-MAR-2000 (first entry)  
 XX  
 DE cDNA clone HTGED19RB related to human interleukin-22 CDNA sequence.  
 XX  
 KW Human; interleukin-22; IL-22; IL-21; immune system disorder;  
 KW immune cell chemotaxis; haematopoietic cell disorder;  
 KW haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;  
 KW respiratory activity; organ rejection; graft-versus-host disease; GVHD;  
 KW inflammation; hyperproliferative disorder; tissue regeneration;  
 KW embryonic stem cell differentiation; embryonic stem cell proliferation;  
 KW haematopoietic lineage; allergic asthma; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9961617-A1.  
 XX  
 PD 02-DEC-1999.  
 XX  
 PF 27-MAY-1999; 99WO-US11644.  
 XX  
 PR 29-MAY-1998; 98US-0087340.  
 PR 10-SEP-1998; 98US-0099805.  
 PR 30-APR-1999; 99US-0131965.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Ruben SM, Ebner R;

XX  
 DR WPI: 2000-072622/06.  
 XX  
 PT Novel polynucleotides used to develop products for treating e.g. immune  
 XX disorders, blood disorders, autoimmune disorders, allergies,  
 PT inflammation, hyperproliferative disorders or infections -  
 XX  
 PS Disclosure; Page 168; 170pp; English.  
 XX  
 CC The present sequence represents a cDNA clone that is related to cDNA  
 CC encoding human interleukin-22 (IL-22) protein. The specification also  
 CC describes IL-21 polynucleotides and polypeptides. The IL-21  
 CC polynucleotide was isolated from a cDNA library of apoptotic T-cells.  
 CC IL-21 and IL-22 may be useful in treating deficiencies or disorders of  
 CC the immune system, by activating or inhibiting the proliferation,  
 CC differentiation, or mobilization (chemotaxis) of immune cells, treating  
 CC or detecting deficiencies or disorders of haematopoietic cells, to  
 CC modulate haemostatic or thrombolytic activity, in treating or detecting  
 CC autoimmune disorders, treating asthma (particularly allergic asthma)  
 CC or other respiratory problems, to treat and/or prevent organ rejection  
 CC or graft-versus-host disease (GVHD), to modulate inflammation, to  
 CC treat or detect hyperproliferative disorders, to treat or detect  
 CC infectious agents, to differentiate, proliferate and attract cells,  
 CC leading to the regeneration of tissues, IL-21 and IL-22 may also  
 CC increase or decrease the differentiation or proliferation of embryonic  
 CC stem cells and haematopoietic lineage, may be used to modulate mammalian  
 CC characteristics.  
 CC  
 XX Sequence 332 BP; 55 A; 112 C; 97 G; 60 T; 8 other;  
 SQ  
 Query Match 20.0%; Score 141.2; DB 21; Length 332;  
 Best Local Similarity 88.9%; Pred. No. 8.2e-19;  
 Matches 176; Conservative 0; Mismatches 15; Indels 7; Gaps 2;  
 QY 9 ttgacacgattgagacccctatccacagaagctgcttcgagctgctgtgacagag 68  
 Db 1 ttgacacgattgagacccctatccacagaagctgcttcgagctgctgtgacagag 60  
 QY 69 gctgtatcagtcacagagcggcgccgagacagctgcttcaactcgtgctgctgctc 128  
 Db 61 gctgtatcagtcacagagcggcgccgagacagctgcttcaactcgtgctgctgctc 120  
 QY 129 agagcctg-citgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 181  
 Db 121 agagcctgagctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 180  
 QY 182 cccacacctggggccttt 199  
 Db 181 cccacacctggggccttt 198  
 RESULT 7  
 O50573 050573 standard; cDNA to mRNA; 1453 BP.  
 ID O50573  
 XX  
 AC O50573;  
 XX  
 DT 24-MAY-1994 (first entry)  
 XX  
 DE Asparaginylendopeptidase clone 104.  
 XX  
 KW Asparaginylendopeptidase; Canavalia ensiformis; seed;  
 KW L-asparagine; primer; PCR; protein fragmentation;  
 KW peptide synthesis; ss.  
 XX  
 OS Canavalia ensiformis.  
 XX  
 PN Key Location/Qualifiers  
 FT CDS 3..1094  
 FT /\*tag= a  
 PN JP05276960-A.



CC	The invention relates to 40 human secreted proteins (Y94881-Y95020),
CC	and cDNA sequences encoding them (A23423-A23462). The secreted proteins
CC	of the invention include those that are thought to be only partially
CC	secreted, i.e., transmembrane proteins. The proteins of the invention may
CC	exhibit one or more activities selected from the following: cytokine
CC	activity; cell proliferation; differentiation; immune modulation;
CC	haematopoiesis regulation; tissue growth activity; activin/inhibin
CC	activity; chemotactic/chemokinetic activity; haemostatic and
CC	thrombolytic activity; anti-inflammatory activity; and tumour inhibition
CC	activity. The proteins may be administered to patients as vaccines, and
CC	the nucleotides may be used as part of a gene therapy regime, diseases or
CC	conditions that may be treated using the proteins or nucleotides of the
CC	invention include autoimmune diseases; genetic disorders; haemophilia;
CC	cardiovascular diseases; cancer; bacterial, fungal and viral infections,
CC	especially HIV; multiple sclerosis; Rheumatoid arthritis; pulmonary
CC	inflammation; Guillain-Barre syndrome; insulin dependent diabetes
CC	mellitus; and allergic reactions such as asthma and anaemia. They may
CC	also be used for treating wounds, burns, ulcers, osteoporosis,
CC	osteoarthritis, periodontal diseases, Alzheimer's disease, Parkinson's
CC	disease, Huntington's disease and amyotrophic lateral sclerosis (ALS).
CC	Proteins with activin/inhibin activity may additionally be useful as
CC	contraceptives. Nucleic acid sequences of the invention may be used in
CC	chromosome mapping, and as a source of diagnostic primers and probes.
CC	The present sequence represents CDNA encoding one of the 40 proteins of
CC	the invention.
XX	
SQ	Sequence 2010 BP; 421 A; 661 C; 592 G; 336 T; 0 other:
Query Match	9.3%; Score 65.8; DB 21; Length 2010;
Best Local Similarity	60.2%; Pred. No. 0.0023;
Matches 109; Conservative	0; Mismatches 72; Indels
Gaps	0;
OY	525 tgcctgtcccggttccttaccatcactgagtgcccggaggctgctct 584 
Db	1808 ttccgagtcacagccccccctgagcccaacctctccaaggagccctccaggaagccct 1867
OY	585 tcccaacctcttggaagtacccctggtttcttaacaattattaagtgtacgttat 644 
Db	1868 tcctcgactcctgagccacagcgacgcctaataaaaactcgtgctttaaaaaaaaa 1927
OY	645 ttaaactgtatcacacaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 704 
Db	1928 aa 1987
OY	705 a 705
Db	1988 a 1988
RESULT	11 .
ID	XI16675 XI16675 standard; DNA; 1554 BP.
XX	
AC	XI16675;
DT	29-APR-1999 (first entry)
DE	Xenopus WA545 protein encoding DNA.
XX	
KW	Xenopus; WA545 protein; TGF-beta; transforming growth factor beta; proliferation; differentiation; mesodermal tissue; neural; muscle; bone; cartilage; connective tissue; wound healing; gene therapy; ss.
OS	Xenopus sp.
XX	
FH	Key Location/Qualifiers
FT	CDS 55..1119
FT	/*tag= a
FT	/product= "WA545 protein"
FT	sig_peptide 55..774
FT	/*tag= b
TT	mat_peptide 775..1116



```
FT XX /*tag= a
PN XX WO9841539-A2.
XX XX
PD XX 24-SEP-1998.
XX XX
PF XX 19-MAR-1998; 98WMO-US05474.
XX XX
PR XX 18-MAR-1998; 98US-0040963.
XX XX 19-MAR-1997; 97US-0820493.
XX XX
PA XX (GENEY ) GENETICS INST INC.
PI XX Agostino MJ, Jacobs K, Lavallie ER, McCoy JM, Merberg D,
XX XX Racie LA, Spaulding V, Treacy M;
XX XX WPI; 1998-521163/44.
DR XX P-PSDB; W79096.
XX XX
XX XX New polynucleotide(s) encoding secreted human proteins - derived
PT PT from human foetal kidney, adult testes and adult or foetal brain
PT PT cDNA libraries
XX XX
XX XX Claim 36; Page 88-91; 112pp; English.
XX XX
CC CC This full-length cDNA clone, designated fe366_1, codes for a novel
CC secreted human protein (see W79096). It was isolated from a human
CC adult brain cDNA library using methods which are selective for
CC cDNAs encoding secreted proteins, or was identified as encoding
CC a secreted or transmembrane protein on the basis of computer
CC analysis of the encoding protein. The nucleotide sequence shows
CC homology to some database sequences, and may contain a CNA repeat
CC and/or Alu repetitive element. The invention provides cDNA clones
CC (see V61477-87) from human foetal kidney, adult testis, and adult
CC or foetal brain cDNA libraries that code for secreted proteins
CC (see W79087-97). These clones are deposited as ATCC 98364. The
CC polynucleotides and proteins are predicted to have useful
CC biological activities which would make them suitable for treating,
CC preventing or ameliorating medical conditions in humans and
CC animals, although no supporting data is given. Suggested
CC activities include nutritional, immune stimulating (e.g. as
CC vaccines) or suppressing, haematopoiesis regulating, tissue growth,
CC activin/inhibin, chemotactic/chemokinetic, haemostatic and
CC thrombolytic, receptor/ligand, antiinflammatory, cadherin/tumour
CC invasion suppressor and tumour inhibition activities. The
CC polynucleotides are also stated to be useful for gene therapy.
XX XX
SQ Sequence 4237 BP; 1330 A; 778 C; 784 G; 1338 T; 7 other;
```

```
KW KW Human: secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX XX
XX XX Homo sapiens.
XX XX
XX XX WO9839448-A2.
XX XX
XX XX 11-SEP-1998.
XX XX
XX XX 06-MAR-1998; 98WMO-US04493.
XX XX
XX XX 02-OCT-1997; 97US-0061060.
XX XX 07-MAR-1997; 97US-0038621.
XX XX 07-MAR-1997; 97US-0040161.
XX XX 07-MAR-1997; 97US-0040162.
XX XX 07-MAR-1997; 97US-0040163.
XX XX 07-MAR-1997; 97US-0040333.
XX XX 07-MAR-1997; 97US-0040334.
XX XX 07-MAR-1997; 97US-0040336.
XX XX 07-MAR-1997; 97US-0040626.
XX XX 11-APR-1997; 97US-0043311.
XX XX 11-APR-1997; 97US-0043312.
XX XX 11-APR-1997; 97US-0043313.
XX XX 11-APR-1997; 97US-0043314.
XX XX 11-APR-1997; 97US-0043568.
XX XX 11-APR-1997; 97US-0043569.
XX XX 11-APR-1997; 97US-0043576.
XX XX 11-APR-1997; 97US-0043578.
XX XX 11-APR-1997; 97US-0043580.
XX XX 11-APR-1997; 97US-0043669.
XX XX 11-APR-1997; 97US-0043670.
XX XX 11-APR-1997; 97US-0043671.
XX XX 11-APR-1997; 97US-0043672.
XX XX 11-APR-1997; 97US-0043674.
XX XX 23-MAY-1997; 97US-0047582.
XX XX 23-MAY-1997; 97US-0047581.
XX XX 23-MAY-1997; 97US-0047583.
XX XX 23-MAY-1997; 97US-0047584.
XX XX 23-MAY-1997; 97US-0047585.
XX XX 23-MAY-1997; 97US-0047586.
XX XX 23-MAY-1997; 97US-0047587.
XX XX 23-MAY-1997; 97US-0047588.
XX XX 23-MAY-1997; 97US-0047589.
XX XX 23-MAY-1997; 97US-0047590.
XX XX 23-MAY-1997; 97US-0047592.
XX XX 23-MAY-1997; 97US-0047593.
XX XX 23-MAY-1997; 97US-0047594.
XX XX 23-MAY-1997; 97US-0047595.
XX XX 23-MAY-1997; 97US-0047596.
XX XX 23-MAY-1997; 97US-0047597.
XX XX 23-MAY-1997; 97US-0047598.
XX XX 23-MAY-1997; 97US-0047599.
XX XX 23-MAY-1997; 97US-0047600.
XX XX 23-MAY-1997; 97US-0047601.
XX XX 23-MAY-1997; 97US-0047612.
XX XX 23-MAY-1997; 97US-0047613.
XX XX 23-MAY-1997; 97US-0047614.
XX XX 23-MAY-1997; 97US-0047615.
XX XX 23-MAY-1997; 97US-0047617.
XX XX 23-MAY-1997; 97US-0047618.
XX XX 23-MAY-1997; 97US-0047632.
XX XX 23-MAY-1997; 97US-0047633.
```





PR	23-MAY-1997	97US-0047492
PR	23-MAY-1997	97US-0047501
PR	23-MAY-1997	97US-0047502
PR	23-MAY-1997	97US-0047503
PR	23-MAY-1997	97US-0047581
PR	23-MAY-1997	97US-0047582
PR	23-MAY-1997	97US-0047583
PR	23-MAY-1997	97US-0047584
PR	23-MAY-1997	97US-0047585
PR	23-MAY-1997	97US-0047586
PR	23-MAY-1997	97US-0047587
PR	23-MAY-1997	97US-0047588
PR	23-MAY-1997	97US-0047590
PR	23-MAY-1997	97US-0047592
PR	23-MAY-1997	97US-0047593
PR	23-MAY-1997	97US-0047594
PR	23-MAY-1997	97US-0047595
PR	23-MAY-1997	97US-0047596
PR	23-MAY-1997	97US-0047597
PR	23-MAY-1997	97US-0047598
PR	23-MAY-1997	97US-0047599
PR	23-MAY-1997	97US-0047600
PR	23-MAY-1997	97US-0047601
PR	23-MAY-1997	97US-0047612
PR	23-MAY-1997	97US-0047613
PR	23-MAY-1997	97US-0047614
PR	23-MAY-1997	97US-0047615
PR	23-MAY-1997	97US-0047617
PR	23-MAY-1997	97US-0047618
PR	23-MAY-1997	97US-0047632
PR	23-MAY-1997	97US-0047633
PR	23-MAY-1997	97US-0047634
PR	23-MAY-1997	97US-0048664
PR	06-JUN-1997	97US-0048974
PR	06-JUN-1997	97US-0048975
PR	13-JUN-1997	97US-0049610
PR	08-JUL-1997	97US-0051326
PR	16-JUL-1997	97US-0052874
PR	18-AUG-1997	97US-0055724
PR	22-AUG-1997	97US-0056630
PR	22-AUG-1997	97US-0056631
PR	22-AUG-1997	97US-0056632
PR	22-AUG-1997	97US-0056633
PR	22-AUG-1997	97US-0056636
PR	22-AUG-1997	97US-0056637
PR	22-AUG-1997	97US-0056642
PR	22-AUG-1997	97US-0056643
PR	22-AUG-1997	97US-0056644
PR	22-AUG-1997	97US-0056645
PR	22-AUG-1997	97US-0056646
PR	22-AUG-1997	97US-0056647
PR	22-AUG-1997	97US-0056676
PR	22-AUG-1997	97US-0056677
PR	22-AUG-1997	97US-0056678
PR	22-AUG-1997	97US-0056679
PR	22-AUG-1997	97US-0056680
PR	22-AUG-1997	97US-0056681
PR	22-AUG-1997	97US-0056682
PR	22-AUG-1997	97US-0056683
PR	22-AUG-1997	97US-0056684
PR	22-AUG-1997	97US-0056685
PR	22-AUG-1997	97US-0056686
PR	22-AUG-1997	97US-0056687
PR	22-AUG-1997	97US-0056688
PR	22-AUG-1997	97US-0056689
PR	22-AUG-1997	97US-0056690
PR	22-AUG-1997	97US-0056691
PR	22-AUG-1997	97US-0056692
PR	22-AUG-1997	97US-0056693
PR	22-AUG-1997	97US-0056694
PR	22-AUG-1997	97US-0056695
PR	22-AUG-1997	97US-0056696
PR	22-AUG-1997	97US-0056697
PR	22-AUG-1997	97US-0056698
PR	22-AUG-1997	97US-0056699
PR	22-AUG-1997	97US-0056700
PR	22-AUG-1997	97US-0056701
PR	05-SEP-1997	97US-0057650

```

PR      05-SEP-1997;       97US-0057669.
PR      05-SEP-1997;       97US-0057761.
PR      12-SEP-1997;       97US-0058785.
XX
PA      (HUMA-) HUMAN GENOME SCI INC.
PI      Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA,
PI      Feng P, Fertile AM, Fischer CL, Florence KA, Greene JM, Hu JS,
PI      Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA,
PI      Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
XX
DR      WPI: 1998-506364/43.
DR      P-FSDB: W74903.
XX
PT      New isolated human genes and the secreted polypeptide(s) they encode
PT      - useful for diagnosis and treatment of e.g. cancers, neurological
PT      disorders, immune diseases, inflammation or blood disorders
XX
PM      Claim 1: Page 415-416; 721pp; English.
XX
PS      This sequence represents a nucleic acid molecule designated Gene 176 from
PS      the human cDNA clone HMPBR48 (deposited as clone ATCC 97904 and ATCC
PS      209050) which encodes a secreted human protein. The gene can be used to
PS      generate fusion proteins by linking to the gene to a human immunoglobulin
PS      Fc portion (e.g. V59502) for increasing the stability of the fused
PS      protein as compared to the human protein only.
PS      The invention relates to 186 novel genes and their fragments (nucleic
PS      acid sequences: V59511-V59812; amino acid sequences W74731-W75026) which
PS      are useful for preventing, treating or ameliorating medical conditions
PS      e.g. by protein or gene therapy. Also, pathological conditions can be
PS      diagnosed by determining the amount of the new polypeptides in a sample
PS      or by determining the presence of mutations in the new polynucleotides.
PS      Specific uses are described for each of the 186 polynucleotides, based on
PS      which tissues they are most highly expressed in (see V59511 for described
PS      uses).
XX
SQ      Sequence 1212 BP; 363 A; 241 C; 307 G; 300 T; 1 other;
XX
Query Match          8.7%: Score 61; DB 19; Length 1212;
Match Local Similarity 64.5%; Pred. No. 0.0017;
Matches 91; Conservative 0; Mismatches 50; Indels 0; Gaps 0
XX
QY      565 gggccgcgagcgctgccttcgccacccttgtagagtcacccctgttctaacaatt 624
Db      1064 gccatcccttggtatatcatatgaacttttaaacattcttcttgtaataaat 1123
QY      625 attaaagtacgltgatattatcaaacgtatgacacaaaaaaaiaaaaaaaaaa 684
Db      1124 aaataagaataaaggctagttcttatgtaalgcaaaaaaaiaaaaaaaaaa 1183
QY      685 aaaaaaaaaaaaaaaaaaaaaa 705
Db      1184 aaaaaaaaaaaaaaaaaaaaaa 1204

```



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2001, 18:18:12 : Search time 2862.16 Seconds  
(without alignments)  
1260.591 Million cell updates/sec

Title: US-09-320-713-1  
Perfect score: 705  
Sequence: 1 gscacgagtgacacgagatg.....aaaaaaaaaaaaaaaaaaaaa 705

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 118133 seqs, 2558875100 residues  
Total number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:\*  
1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_om:\*  
4: gb\_ov:\*  
5: gb\_ph:\*  
6: gb\_pl1:\*  
7: gb\_pl2:\*  
8: gb\_pr1:\*  
9: gb\_pr2:\*  
10: gb\_pr3:\*  
11: em\_fun:\*  
12: em\_hum1:\*  
13: em\_hum2:\*  
14: em\_in:\*  
15: em\_om:\*  
16: em\_ov:\*  
17: em\_ph:\*  
18: em\_pat:\*  
19: em\_pl:\*  
20: em\_ro:\*  
21: em\_sts:\*  
22: em\_sy:\*  
23: em\_un:\*  
24: em\_vl:\*  
25: em\_vl:\*  
26: gb\_htg1:\*  
27: gb\_htg2:\*  
28: gb\_in1:\*  
29: gb\_in2:\*  
30: em\_ba1:\*  
31: em\_ba2:\*  
32: em\_hum3:\*  
33: em\_hum4:\*  
34: gb\_pr4:\*  
35: gb\_htg3:\*  
36: gb\_htg4:\*  
37: gb\_htg5:\*  
38: gb\_htg6:\*  
39: gb\_htg7:\*  
40: em\_htg1:\*  
41: em\_htg2:\*  
42: em\_htg3:\*  
43: em\_hum5:\*

44: gb\_pl3:\*  
45: gb\_pr5:\*  
46: gb\_htg8:\*  
47: gb\_htg9:\*  
48: gb\_htg10:\*  
49: gb\_htg11:\*  
50: gb\_htg12:\*  
51: gb\_htg13:\*  
52: gb\_htg14:\*  
53: gb\_in3:\*  
54: gb\_htg15:\*  
55: gb\_htg16:\*  
56: gb\_htg17:\*  
57: em\_htg4:\*  
58: em\_htg5:\*  
59: em\_htg6:\*  
60: em\_htg7:\*  
61: em\_hum6:\*  
62: gb\_htg18:\*  
63: gb\_htg19:\*  
64: gb\_ba3:\*  
65: em\_htg8:\*  
66: em\_htg9:\*  
67: em\_htg10:\*  
68: gb\_pr6:\*  
69: gb\_pr7:\*  
70: gb\_htg20:\*  
71: gb\_htg21:\*  
72: gb\_htg22:\*  
73: gb\_htg23:\*  
74: gb\_ro:\*  
75: gb\_sts1:\*  
76: gb\_sts2:\*  
77: gb\_sy:\*  
78: gb\_un:\*  
79: gb\_vl1:\*  
80: gb\_vl2:\*  
81: gb\_pat1:\*  
82: gb\_pat2:\*  
83: em\_htg0:\*  
84: gb\_htg24:\*  
85: gb\_pr8:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	650.4	92.3	157090	47	AC022554 Homo sapi
2	640.4	90.8	1047	34	AF152099 Homo sapi
3	64.2	9.1	1315	74	AF082526 Mus muscu
4	61.8	8.8	989	7	CPCRM7
5	61.4	8.7	1554	4	AF065135
6	61.2	8.7	1221	69	HSMB01506
7	60.8	8.6	3469	69	HSMB00438
8	60.4	8.6	6410	34	AF230496
9	60.4	8.6	1478	34	AK025967
10	60.4	8.6	1678	69	HSMB02263
11	60.2	8.5	2358	3	AF254119
12	60	8.5	1419	7	CCCHSMR
13	60	8.5	3336	69	HSMB01246
14	59.8	8.5	1809	7	ATPBR2
15	59.8	8.5	4173	69	HSMB01960
16	59.6	8.5	7218	81	166494
17	59.4	8.4	1523	10	AF116614
18	59.4	8.4	49473	52	AC061987
19	59.2	8.4	2156	34	AK025435
20	59	8.4	1560	34	AK026057
21	59	8.4	1739	44	PC042442

22	59	8.4	1826	7	AF196975	AF196975	Pneumocys
23	58.8	8.3	1056	74	MUSUNKNB	L04849	Mouse (Clon
24	58.8	8.3	1405	69	HSM800974	AL117463	Homo sapi
25	58.8	8.3	2341	10	AF119857	AF119857	Homo sapi
26	58.8	8.3	2345	74	MMU87620	U87620	Mus musculi
27	58.8	8.3	4212	69	HSM800772	AL110269	Homo sapi
28	58.6	8.3	1740	10	AF069506	AF069506	Homo sapi
29	58.6	8.3	1845	69	HSM800467	AL050172	Homo sapi
30	58.6	8.3	110000	84	PFMAL13P2_0	AL049185	Plasmodiu
31	58.4	8.3	617	85	IROEST087	AL357195	Homo sapi
32	58.4	8.3	2079	8	AB047878	AB047878	Macaeca fa
33	58.4	8.3	2665	34	AF260566	AF260566	Homo sapi
34	58.2	8.3	1602	81	AR000496	AR000496	Sequence
35	58.2	8.3	1699	69	HSU39656	U39656	Human MAP k
36	58.2	8.3	1798	74	AF246457	AF246457	Rattus no
37	58.2	8.3	1838	74	AF246458	AF246458	Rattus no
38	58.2	8.3	1897	81	IS2013	IS2013	Sequence 1
39	58	8.2	1288	53	TBBDPMSTN	254162	T.Brucel br
40	58	8.2	2344	34	AK024855	AK024855	Homo sapi
41	58	8.2	2827	69	HSM801964	AL137660	Homo sapi
42	58	8.2	4632	69	HSU72621	U72621	Homo sapien
43	57.8	8.2	522	34	AK026629	AK026629	Homo sapi
44	57.8	8.2	721	34	AK026526	AK026526	Homo sapi
45	57.8	8.2	1162	7	AF275315	AF275315	Lotus jap

## ALIGNMENTS

RESULT 1  
AC022554  
LOCUS  
DEFINITION Homo sapiens clone RP11-21B21, WORKING DRAFT SEQUENCE, 35 unordered  
pieces.  
AC022554  
AC022554.2 GI:7637249  
HTG: HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE  
ORGANISM  
Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
1 (bases 1 to 157090)  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
Bliren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Becker,I., Beda,F.,  
Boguslavskiy,L., Bouhgalter,B., Brown,A., Burkett,G., Castle,A.,  
Choeppel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,  
DeRellano,K., Dewar,K., Domino,M., Doyle,M., Fenesstor,J.,  
Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,  
Gardyna,S., Grant,G., Hagos,B., Heatford,A., Horton,L.,  
Howard,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,  
Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,  
Macdonald,P., Marquis,N., McEwan,P., McGuirk,A., McKernan,K.,  
McNeeters,R., Meldrum,J., Menus,L., Morrow,J., Naylor,J.,  
Norman,C.H., O'Connor,T., O'Donnell,P., Oliver,T.M., Peterson,K.,  
Pierre,N., Pisanl,C., Pollara,V., Raymond,C., Riley,R., Rotman,D.,  
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,  
Stojanovic,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,  
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,  
Zimmer,A. and Zody,M.  
2 (bases 1 to 157090)  
Unpublished

TITLE  
JOURNAL  
COMMENT  
Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 22, 2000 this sequence version replaced g1:6910831.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR

Web site: <http://www.seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L2245  
Center clone name: 21\_B\_21  
----- Summary Statistics  
Sequencing vector: M13; M7815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 138238 bases at least Q40  
Consensus quality: 147017 bases at least Q30  
Consensus quality: 150850 bases at least Q20  
Insert size: 170000; agarose-fp  
Insert size: 153690; sum-of-contigs  
Quality coverage: 3.1 in Q20 bases; agarose-fp  
Quality coverage: 3.5 in Q20 bases; sum-of-contigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 35 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1  
1661: contig of 1661 bp in length  
1662 1761: gap of 100 bp  
1762 3078: contig of 1317 bp in length  
3079 3178: gap of 100 bp  
3179 4774: contig of 1596 bp in length  
4775 4874: gap of 100 bp  
4875 7439: contig of 2565 bp in length  
7440 7539: gap of 100 bp  
7540 8885: contig of 1346 bp in length  
8886 8985: gap of 100 bp  
8986 11624: contig of 2639 bp in length  
11625 11724: gap of 100 bp  
11725 14597: contig of 2873 bp in length  
14598 14697: gap of 100 bp  
14698 16860: contig of 2163 bp in length  
16861 16960: gap of 100 bp  
16961 19655: contig of 2695 bp in length  
19656 19755: gap of 100 bp  
19756 21926: contig of 2171 bp in length  
21927 22026: gap of 100 bp  
22027 23792: contig of 1766 bp in length  
23793 23892: gap of 100 bp  
23893 27939: contig of 4047 bp in length  
27940 28039: gap of 100 bp  
28040 30269: contig of 2230 bp in length  
30270 30369: gap of 100 bp  
30370 33004: contig of 2635 bp in length  
33005 33104: gap of 100 bp  
33105 36260: contig of 3156 bp in length  
36261 36360: gap of 100 bp  
36361 39762: contig of 3402 bp in length  
39763 39862: gap of 100 bp  
39863 43798: contig of 3936 bp in length  
43799 43898: gap of 100 bp  
43899 48559: contig of 4661 bp in length  
48560 48659: gap of 100 bp  
48660 52026: contig of 3367 bp in length  
52027 52126: gap of 100 bp  
52127 55169: contig of 3043 bp in length  
55170 55268: gap of 100 bp  
55270 58796: contig of 3527 bp in length  
58797 58896: gap of 100 bp  
58897 62413: contig of 3517 bp in length  
62414 62513: gap of 100 bp  
62514 67190: contig of 4677 bp in length  
67191 67290: gap of 100 bp  
67291 71331: contig of 4441 bp in length  
71332 71431: gap of 100 bp



BASE COUNT	ORIGIN	CDS
177 a	381 c	280 g
209 t		
		/organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="2" /map="2q31" /clonene="DNA62377" 50. .643 /note="IL-17C" /codon_start=1 /product="interleukin 17C" /protein_id="AA28105.1" /db_xref="GI:6759915" /translation="MTLLPGLFTWLHTCLAHHPDLRGPHSHGTPHCYSAAELPL GQAPRLVLAGAKWQALPAVALYSLEAASHARHNEPSATQCPVLRPEVLTADTH GASPSPIRYVDDEEDRYPOKLPAFECIRCCIDARQRETAALNSVRLQSLTLVLR RPSRDSGSLTPCAFAPHFHFIVHPGTCVLRVSV"

[illegible]

RESULT	3
AF082526	
LOCUS	AF082526
DEFINITION	1315 bp mRNA
ACCESSION	AF082526
VERSION	AF082526.1
KEYWORDS	GI:3549604
SOURCE	house mouse.
ORGANISM	Mus musculus
	ROD
	107-SEP-1998
	complete cds.

REFERENCE  
AUTHORS  
Eukaryota; Metazoa; Chordata; Crustacea; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus  
1 (bases 1 to 1315)  
Schaeffer, H.J., Catling, A.D., Edlen, S.T., Collier, L.S., Krauss, A.  
and Weber, M. J.  
TITLE  
MPL: A MKK binding partner that enhances enzymatic activation of  
the MAP kinase cascade  
JOURNAL  
Unpublished  
REFERENCE  
AUTHORS  
2 (bases 1 to 1315)  
Schaeffer, H.J., Catling, A.D. and Weber, M. J.  
TITLE  
Direct Submission  
JOURNAL  
Submitted (05-AUG-1998) Microbiology, University of Virginia, JPA  
1300, Charlottesville, VA 22908, USA  
FEATURES  
Location/Qualifiers

```

source
  1. .1315
  /organism="Mus musculus"
  /db_xref="taxon:10090"
gene
  1. .1315
  /gene="Mpl"
  147. .521
  /gene="Mpl"
  /codon_start=1
  /product="MEK binding partner 1"
  /protein_id="AAC34591.1"
  /db_xref="GI:3549605"
  /translation="MADDIKRLFLYKKRIPSVEGLHAIYVSDRDGVPTIKVANDSPAEHA
  LRREFLSTPALATDQSKLGLSKKSKSTICTYNTYQVQFNRPLPVLSFIASSSANTGL
  IVLSKEKLAFLFEELIVEVS"
BASE COUNT      395 a      273 c      300 g      347 t
ORIGIN

```

Query Match	9.1%:	Score 64.2:	DB 74:	Length 1315:
Best Local Similarity	74.3%:	Pred. No. 0.0094:		
Matches 81:	Conservative	0:	Mismatches 28:	Indels 0:
Gaps				
QY 597	tggaagtaccccctgttcttaacacatatttaagtgacgtgatatttaacatgata	655		
Db 1207	TGACGAGTGGCCATTTTTTTTATCATGTGTATTAATAAATAGATATTTTAAAGGTAAAA	1207		
QY 657	acacacaaa	705		
Db 1267	AAA	1315		

FEATURES	Source
JOURNAL	development in Xenopus
MEDLINE	Development 126 (7), 1467-1482 (1999)
REFERENCE	99169090
AUTHORS	2 (bases 1 to 1554)
TITLE	Sun, B.-I.
JOURNAL	Direct Submission
FEATURES	Submitted (13-MAY-1998) whitehead Institute, 9 Cambridge Center, Cambridge, MA 02142, USA
Source	Location/Qualifiers
CDS	1..1554
	/organism="Xenopus laevis"
	/db_xref="taxon:8355"
	55..1119
	/function="posterior determination"
	/note="TFEB family member"
	/codon_start=1
	/product="geriere"
	/protein_id="AAD19837.1"
	/db_xref="GI:4405984"
	/translation="MAELWLSLSCMSFLLLTNNSSPYTFQERMLKALGLNTRPNPIA PAPPKSLRIDFEKGINODNFCMEGCGVPNIIVRSYDQGTIAIEPQSLCKRFF LEFLPSAVENKEOLTLGOLTEKFKHNTYGGQFLRLRYRTOLSLKGRDCKMRKILF VTSFRLRHKSLSYFNLTIRVAEDKMKPNKMGMLILEIYASSLAGNSFVCEPLOSSE ITYSLTIVSLDPSNCKIQORAKRTHSSPEPTSNICKRRRLTYDKDVGOMWVIAPRK YMAVYCGECYPTLEMLRGTHNAVLDTLVHSEPEMTPLPCCAPTRKLSPISMLYDND NDNVLRHYEDMVVDEGCK"
BASE COUNT	503 a 340 c 309 g 402 t
ORIGIN	
Query Match	8.7%, Score 61.4; DB 4; Length 1554;
Best Local Similarity	72.1%; Pred. No. 0.032;
Matches	80; Conservative 0; Mismatches 31; Indels 0; Gaps 0;
QY	595 cttagaagtaacccctgttcttaacaaattatttaagtcgctgtatataaactgat 654
Db	1437 CTTGCAAGTACAGCGTGGATTTGGTGAATTAAGTATTAAACCCAAAAA 1496
QY	655 gaacacaaaaa 705
Db	1497 AA 1547
RESULT	6
LOCUS	HSMB01506 mRNA PRI 18-FEB-2000
DEFINITION	Homo sapiens mRNA; CDNA DKFZp34J1618 (from clone DKFZp34J1618); partial cds.
ACCESSION	AL133637
VERSION	AL133637.1 GI:6599285
KEYWORDS	
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS	1 (bases 1 to 1221)
	Ottensmæder, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE	Direct Submission
JOURNAL	Submitted (15-DEC-1999) MIPS, Am Klopfersplitz 18a, D-82152 Martinsried, GERMANY
COMMENT	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Medigenomix (Martinsried/Germany) within the CDNA (DKFZp34J1618) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further information about the clone and the sequencing project is available at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES	Location/Qualifiers
Source	1..1221
	/organism="Homo sapiens"









OY	609	Iglttcttaaacacattattatgaigtactgtatattataactgtagcaccacaaaaaa	668
Db	2253	TATTTAATCTTACTTAAATATAAGTTTCTTTTAAACGAGAAAAA	2312
OY	669	aaaaaaaaaaaaaaaaaaaaaaaaaaaaa	705
Db	2313	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2349
RESULT	12		
CCHSMR			
LOCUS	CCCHSMR	1419 bp	mRNA
DEFINITION		C. chinensis mRNA for chalcone synthase.	PLN
ACCSSION		267988	15-APR-1998
VERSION		267988.1	GI:3059082
KEYWORDS		chalcone synthase.	
SOURCE		Callistephus chinensis.	
ORGANISM		Callistephus chinensis	
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
AUTHORS		eunhyilphytes; Spermatophyta; Magnoliophyta; eudicotyledons;	
REFERENCE		Asteridae; Asterales; Asteraceae; Callistephus.	
AUTHORS		Henkel,J., Wassenegger,M., Sommer,H. and Forkmann,G.	
REFERENCE		Unpublished	
AUTHORS		Henkel,J.	
REFERENCE		Submitted (10-NOV-1995)	
AUTHORS		Muenchen, Lehrstuhl fuer Zierpflanzenbau, Blumenstrasse 16,	
REFERENCE		Freising, Germany	
AUTHORS		Revised by [3]	
REFERENCE		3 (bases 1 to 1419)	
AUTHORS		Henkel,J.	
REFERENCE		Direct Submission	
AUTHORS		Submitted (18-JUL-1997)	
REFERENCE		Muenchen, Lehrstuhl fuer Zierpflanzenbau, Blumenstrasse 16,	
AUTHORS		Freising, Germany	
COMMENT		On Apr 18, 1998 this sequence version replaced gi:1632780.	
FEATURES			
source		location/qualifiers	
		1..1419	
		/organism="Callistephus chinensis"	
		/strain="L 01 (Institute collection)"	
		/db_xref="taxon:13379"	
		/dev_stage="buds and flowers"	
		/tissue_type="petals"	
		33..1229	
		/EC_number="2.3.1.74"	
		/codon_start=1	
		/product="chalcone synthase"	
		/protein_id="CAA91930.1"	
		/db_xref="GI:3059083"	
		/db_xref="SWISS-PROT:P48365"	
		/translation="MASTIDIAIREKORROGPATITAIIGTAPNSCVQADYPDYEE	
		RITSEHNVDLKEFKRMCDKSMIRKRMRHTTEYLKENPSLCYMASLSIARODVVV	
		VEVRLKEAKTKAIKWGGPKSKITPHIFCTGTGVMPGADYOLTKILGRPSVKERF	
		MGMOQCFCAGTGYVLRIANKDLAENNKGARLVAVCEITAYVRPGNDHLDSLGCALE	
		GDGAAYIVGADPDLTTERLPFLPEFMISAQOTILPSEGALDGLREVGITLPHLLDVGG	
		LISNIKKALTOASPIGITDWNISIPWLAHPGPAIIIDOVETLKIKLFEKKRATRHHVI	
		SEYNMSACVLFITIDEKKRSADGAAITTEGIDMGVLEFGTLIVEIVLHSLPT	
		TMAIAT"	
		polya_signal	
		1298..1303	
		polya_signal	
		1332..1337	
		polya_site	
		1357	
BASE COUNT		409 a	320 c 333 g 357 t
ORIGIN			
Query Match		8.5%; Score 60; DB 7; Length 1419;	
Best Local Similarity		75.0%; Pred. No. 0.059;	
Matches		75; Conservative 0; Mismatches 25; Indels 0; Gaps 0;	
OY	606	cccgtttctttaacaacattattatgaigtactgtatattataactgtagcaccacaaaaa	665

[illegible]





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2001, 19:08:07 ; Search time 1413.06 Seconds  
(without alignments)  
3496.156 Million cell updates/sec

Title: US-09-320-713-1  
Perfect score: 705  
Sequence: 1 ggcacgagtgacacgagtg.....aaaaaaaaaaaaaaaaaaaaa 705

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 7991742 seqs, 3503743858 residues

Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_est3:\*  
4: gb\_est4:\*  
5: gb\_est5:\*  
6: gb\_est6:\*  
7: gb\_est7:\*  
8: gb\_est8:\*  
9: gb\_est9:\*  
10: gb\_est10:\*  
11: gb\_est11:\*  
12: gb\_est12:\*  
13: gb\_est13:\*  
14: gb\_est14:\*  
15: gb\_est15:\*  
16: gb\_est16:\*  
17: gb\_est17:\*  
18: gb\_est18:\*  
19: gb\_est19:\*  
20: gb\_est20:\*  
21: gb\_est21:\*  
22: gb\_est22:\*  
23: gb\_est23:\*  
24: gb\_est24:\*  
25: gb\_est25:\*  
26: gb\_est26:\*  
27: gb\_est27:\*  
28: gb\_est28:\*  
29: gb\_est29:\*  
30: gb\_est30:\*  
31: gb\_est31:\*  
32: gb\_est32:\*  
33: gb\_est33:\*  
34: gb\_est34:\*  
35: gb\_est35:\*  
36: gb\_est36:\*  
37: gb\_est37:\*  
38: gb\_est38:\*  
39: gb\_est39:\*  
40: gb\_est40:\*  
41: em\_estda:\*  
42: em\_estfun:\*  
43: em\_esthum1:\*

44: em\_esthum2:\*  
45: em\_esthum3:\*  
46: em\_esthum4:\*  
47: em\_esthum5:\*  
48: em\_esthum6:\*  
49: em\_esthum7:\*  
50: em\_esthum8:\*  
51: em\_esthum9:\*  
52: em\_esthum10:\*  
53: em\_esthum11:\*  
54: em\_esthum12:\*  
55: em\_esthum13:\*  
56: em\_esthum14:\*  
57: em\_esthum15:\*  
58: em\_esthum16:\*  
59: em\_esthum17:\*  
60: em\_esthum18:\*  
61: em\_esthum19:\*  
62: em\_esthum20:\*  
63: em\_estin1:\*  
64: em\_estin2:\*  
65: em\_estin3:\*  
66: em\_estin4:\*  
67: em\_estov1:\*  
68: em\_estov2:\*  
69: em\_estp11:\*  
70: em\_estp12:\*  
71: em\_estp13:\*  
72: em\_estp14:\*  
73: em\_estp15:\*  
74: em\_estro1:\*  
75: em\_estro2:\*  
76: em\_estro3:\*  
77: em\_estro4:\*  
78: em\_estro5:\*  
79: em\_estro6:\*  
80: em\_estro7:\*  
81: em\_estro8:\*  
82: em\_estro9:\*  
83: em\_estro10:\*  
84: em\_estro11:\*  
85: em\_estro12:\*  
86: em\_estro13:\*  
87: gb\_esta1:\*  
88: gb\_esta2:\*  
89: gb\_esta3:\*  
90: gb\_esta4:\*  
91: gb\_esta5:\*  
92: gb\_esta6:\*  
93: gb\_esta7:\*  
94: gb\_esta8:\*  
95: gb\_esta9:\*  
96: gb\_esta10:\*  
97: gb\_esta11:\*  
98: gb\_esta12:\*  
99: gb\_esta13:\*  
100: gb\_esta14:\*  
101: gb\_esta15:\*  
102: gb\_esta16:\*  
103: gb\_esta17:\*  
104: gb\_esta18:\*  
105: gb\_esta19:\*  
106: gb\_esta20:\*  
107: gb\_esta21:\*  
108: gb\_esta22:\*  
109: gb\_esta23:\*  
110: gb\_esta24:\*  
111: gb\_esta25:\*  
112: em\_esthum21:\*  
113: em\_esthum22:\*  
114: em\_esthum23:\*  
115: em\_estom1:\*  
116: em\_estom2:\*

117: em\_estp16:\*  
 118: em\_estp17:\*  
 119: em\_estp18:\*  
 120: em\_estp19:\*  
 121: em\_estp20:\*  
 122: em\_estp21:\*  
 123: em\_estp22:\*  
 124: em\_estp23:\*  
 125: em\_estp24:\*  
 126: em\_estp25:\*  
 127: em\_estp26:\*  
 128: em\_estp27:\*  
 129: em\_estp28:\*  
 130: em\_estp29:\*  
 131: em\_estp30:\*  
 132: em\_estp31:\*  
 133: em\_estp32:\*  
 134: em\_estp33:\*  
 135: em\_estp34:\*  
 136: em\_estp35:\*  
 137: em\_estp36:\*  
 138: em\_estp37:\*  
 139: em\_estp38:\*  
 140: em\_estp39:\*  
 141: em\_estp40:\*  
 142: em\_estp41:\*  
 143: em\_estp42:\*  
 144: em\_estp43:\*  
 145: em\_estp44:\*  
 146: em\_estp45:\*  
 147: em\_estp46:\*  
 148: em\_estp47:\*  
 149: em\_estp48:\*  
 150: em\_estp49:\*  
 151: em\_estp50:\*  
 152: em\_estp51:\*  
 153: em\_estp52:\*  
 154: em\_estp53:\*  
 155: em\_estp54:\*  
 156: em\_estp55:\*  
 157: em\_estp56:\*  
 158: em\_estp57:\*  
 159: em\_estp58:\*  
 160: em\_estp59:\*  
 161: em\_estp60:\*  
 162: em\_estp61:\*  
 163: em\_estp62:\*  
 164: em\_estp63:\*  
 165: em\_estp64:\*  
 166: em\_estp65:\*  
 167: em\_estp66:\*  
 168: em\_estp67:\*  
 169: em\_estp68:\*  
 170: em\_estp69:\*  
 171: em\_estp70:\*  
 172: em\_estp71:\*  
 173: em\_estp72:\*  
 174: em\_estp73:\*  
 175: em\_estp74:\*  
 176: em\_estp75:\*  
 177: em\_estp76:\*  
 178: em\_estp77:\*  
 179: em\_estp78:\*  
 180: em\_estp79:\*  
 181: em\_estp80:\*  
 182: em\_estp81:\*  
 183: em\_estp82:\*  
 184: em\_estp83:\*  
 185: em\_estp84:\*  
 186: em\_estp85:\*  
 187: em\_estp86:\*  
 188: em\_estp87:\*  
 189: em\_estp88:\*

190: gb\_gss25:\*  
 191: gb\_gss26:\*  
 192: gb\_gss27:\*  
 193: gb\_gss28:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	260	36.9	264	109	BE551188	BE551188 7b55g903.x
2	71	10.1	242	29	AU037868	AU037868 AU037868
3	69.8	9.9	369	39	AM090736	AM090736 XC9502.x
4	69	9.8	366	111	BE723174	BE723174 192592 MA
5	68.8	9.8	359	23	A1673363	A1673363 LW35B08.x
6	67.8	9.6	199	3	AA198893	AA198893 mu05h09.r
7	66.8	9.5	294	22	A1581387	A1581387 co71c06.x
8	66.8	9.5	1058	137	BE905366	BE905366 601499028
9	66	9.4	309	140	C90935	C90935 C90935 Dict
10	65.4	9.3	363	92	AW541819	AW541819 C0149H07-
11	65.2	9.2	394	136	BE808909	BE808909 214216 MA
12	64.8	9.2	225	27	A1961414	A1961414 wt22c07.x
13	64.8	9.2	550	26	A1925744	A1925744 wo34h11.x
14	64.6	9.2	400	90	AM394524	AM394524 sh32b03.y
15	64.6	9.2	1201	190	CNS016DQ	AL106616 Drosophila
16	64.4	9.1	245	4	AA274319	AA274319 TGESTr225
17	64.4	9.1	308	140	C90985	C90985 C90985 Dict
18	64	9.1	263	25	A1818980	A1818980 w189h11.x
19	63.8	9.0	367	87	AM190194	AM190194 x112e08.x
20	63.6	9.0	206	29	AU038875	AU038875 AU038875
21	63.6	9.0	261	25	A1819976	A1819976 w188h11.x
22	63.6	9.0	285	94	AW733216	AW733216 sk71c02.y
23	63.6	9.0	322	27	A1973152	A1973152 wf52c02.x
24	63.6	9.0	349	20	A1475371	A1475371 cl18e01.x
25	63.6	9.0	796	190	CNS0118D	AL099943 Drosophila
26	63.4	9.0	230	22	A1559531	A1559531 tg50b09.x
27	63.4	9.0	348	134	BE058451	BE058451 sn16c06.y
28	63.4	9.0	795	39	AM024360	AM024360 w030c04.x
29	63.4	9.0	822	108	BE420782	BE420782 HMO02.E0
30	63.2	9.0	121	39	AM059474	AM059474 fe14e01.y
31	63.2	9.0	241	22	A1613038	A1613038 ty06h05.x
32	63.2	9.0	337	111	BE667471	BE667471 151857 MA
33	63.2	9.0	371	95	AW786724	AW786724 120133 MA
34	63	8.9	120	22	A1590043	A1590043 tr75c04.x
35	63	8.9	386	18	A1267454	A1267454 ag65b01.x
36	62.8	8.9	134	29	AU053905	AU053905 AU053905
37	62.8	8.9	271	40	AM100582	AM100582 sd57c01.y
38	62.6	8.9	838	106	BE319360	BE319360 NF016E02R
39	62.6	8.9	247	3	AA183005	AA183005 mt78c04.r
40	62.6	8.9	312	4	AA270657	AA270657 va69a07.f
41	62.4	8.9	205	40	AM100861	AM100861 sq62c12.y
42	62.4	8.9	248	18	A1245008	A1245008 qj94b05.x
43	62.4	8.9	294	105	BE234043	BE234043 140614 MA
44	62.2	8.8	224	19	A1395684	A1395684 MA004543
45	62.2	8.8	338	21	A1538637	A1538637 tp70e10.x

## ALIGNMENTS

RESULT 1  
 BE551188/c LOCUS 264 bp mRNA  
 DEFINITION 7b55g903.x1 NC1\_CGAP\_Lu24 Homo sapiens cDNA clone IMAGE:3232180 3',  
 BE551188 mRNA sequence.  
 ACCESSION BE551188  
 VERSION BE551188.1 GI:9792880  
 KEYWORDS EST.  
 SOURCE human.



ORGANISM	Homo sapiens
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

ORGANISM	REFERENCE
<i>Homo sapiens</i>	1 (bases 1 to 264)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	

## БЕЛАТРИЦЕ

## Source

BASE COUNT ORIGIN	66 a	50 c	88 g	60 t
----------------------	------	------	------	------

VERSION

VERSION	AU037868.1	GI:3984621
KEYWORDS	EST.	
SOURCE	Dictyostelium discoideum.	
ORGANISM	Dictyostelium discoideum	

## FEATURES

BASE COUNT	147 a	5 c	1 g	88 t	1 others
ORIGIN					



Query Match	9.88;	Score 68.8;	DB 23;	Length 359;
Best Local Similarity	60.9%;	Pred. No. 5.2e-05;		
Matches 112; Conservative	0;	Mismatches 72;	Indels 0;	Gaps 0;

OY	582	tcttcccaactccttggaagtaaccctgylttctctaacaattatlaaglygtacgta	641
Db	203	CCGGGCGCTCCCGGGGTCGCCGGGTCCTTTTTCGCCCCCCCCCCCCCCGGG	144
OY	522	cactgcctgcgtcccgagcttccttaaccttaacctaactgctcgaagccccgaagctgyc	581
Db	179		144
OY	642	tattaactgtagaacccaataaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa	701
Db	83	AAA	24
OY	702	aaaa 705	
Db	23	 AAAA 20	

RESULT	6
AA198893	
LOCUS	AA198893
DEFINITION	AA198893 199 bp mRNA EST 17-FEB-1997 mu05f09.r1 Soares mouse 3NDMS Mus musculus cDNA clone IMAGE:638561 5', mRNA sequence.

VERSION AA198893.1 GI:1795069

SOURCE	house mouse.
ORGANISM	Mus musculus

REFERENCE  
AUTHORS  
Matta, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T.

Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R. and Waterston, R.

TITLE	The WashU-HHMI Mouse EST Project
JOURNAL	Unpublished (1996)
COMMENT	Contact: Marra M/Mouse EST Project

WASHU-HHMI Mouse EST Project  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314.265.1000

Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: [mouseest@watson.wustl.edu](mailto:mouseest@watson.wustl.edu)

IMAGE Consortium ([infoimage.llnl.gov](http://infoimage.llnl.gov)) for further information.  
MGI:390553  
Seq primer: -28M13 run2 from Amersham

High quality sequence stop: 146.  
Location/Qualifiers  
1 109  
source

```

1: 105
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"

```

```

/clone_1b="Soares mouse 3NbMS"
/sex="male"

```

```

/issue_type="Spleen"
/dev_stage="4 weeks"
/lab_host="DH10B"

```

/note-"Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5].

TGTTACCAATCTGAAGTGGAGCGGCCGCTGTTTTTTTTTTTTTTTTTTT  
3']; double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I

and Eco RI sites of the modified pT7T3 vector. RNA provided by Dr. Bertrand Jordan. Library went through

three rounds of normalization, and was constructed by Bento Soares and M.Fatima Bonaldo."

Query Match	Similarity	6.6%	Score 67.8	DB 3	Length 199
Best Local	Similarity	60.7%	Pred. No. 8.3e-05		
Matches	111	Conservative	0	Mismatches	72
				Indels	0
				Gaps	0
QY	523	ctgtgcctgctgctccgagctctccctatactacgctcagccgccgcagctgctc	582		
Db	17	ccgcttcctgctgctgctgctccgctccgccaacactgtctcttcacatgctctgcagct	76		
QY	583	cttcccaaccctccttggaagtaacccctgtttctttaacaatatttaagtgaagtgtat	642		
Db	77	cttgccccaacagatgacgctgatatattttttcttatctatttttttaataaaaaa	136		
QY	643	tatttaactgtatgacacaaaaaataaaaaaataaaaaaataaaaaaataaaaaa	702		
Db	137	gagtttaaaaaaataaaaaaataaaaaaataaaaaaataaaaaaataaaaaa	196		

QY	703	aaa	705
Db	197	AAA	199

RESULT	7		
AI581387/c			
LOCUS	AI581387	294 bp	mrna
DEFINITION	ct71c06.x1 NCL1.GCsp4 Homo sapiens cDNA clone IMAGE:2183722 3',		EST
	mrna sequence.		

ACCESSION	A1581387
VERSION	A1581387.1
KEYWORDS	EST,
SOURCE	human.

ORGANISM	REFERENCE
<i>Homo sapiens</i>	1 (bases 1 to 294)
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi	
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	

**AUTHORS** NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
**TITLE** National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

**JOURNAL COMMENT**      Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550

Email: Robert\_Strausberg@nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R  
Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 at Washington University

Cione distribution: MCL-GAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINL at:  
[www-bio.lnl.gov/bbrp/image/image.html](http://www-bio.lnl.gov/bbrp/image/image.html)  
 Contact: Joseph C. ... 0.00

Insert length: 1065      Std Error  
Seq primer: -40UP from Gibco  
High quality sequence stop: 249  
POLYVA=NO

COLIN-NO.	Location/Qualifiers
FEATURES	
source	
1. .294	
	/organism="Homo sapiens"

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:218372"
/clone_lib="NCT CGAP Cas4"

```

```

/clone_id= "NCI_C12688"
/tissue_type="poorly differentiated adenocarcinoma with
signet ring cell features"
/lab_host="DH10B"

```

/note="Organ: stomach; Vector: pCMV-SPORT6; Site\_1: SalI  
 Site\_2: NotI; Cloned unidirectionally. Primer: Oligo dT  
 Average insert size 1.69 kb. Life Technologies catalog #

BASE COUNT	79 a	47 c	36 g	132 t
------------	------	------	------	-------



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae

COMMENT Contact: Robert Strausberg, Ph.D.





---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2001, 19:10:02 ; Search time 86.16 Seconds  
(without alignments)  
1318.686 Million cell updates/sec

Title: US-09-320-713-1

Perfect score: 705

Sequence: 1 ggcacgagtgacacgagtg.....aaaaaaaaaaaaaaaaaaaaa 705

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 280836 seqs, 80580151 residues

Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq: \*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq: \*  
3: /cgn2\_6/ptodata/2/ina/6.COMB.seq: \*  
4: /cgn2\_6/ptodata/2/ina/PCtUS.COMB.seq: \*  
5: /cgn2\_6/ptodata/2/ina/Backfile1.seq: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	61.2	8.7	882	2	US-08-909-965C-9
2	59.6	8.5	7218	1	US-08-232-463-14
3	58.2	8.3	1602	1	US-08-530-950-3
4	58.2	8.3	1602	3	US-08-888-429A-3
5	58.2	8.3	1897	1	US-08-184-632-1
6	57.6	8.2	1582	3	US-08-545-186B-10
7	57.6	8.2	1582	3	US-08-545-186B-12
8	57.4	8.1	467	2	US-08-841-349-18
9	57	8.1	2989	5	5378464-1
10	56.6	8.0	1098	3	US-09-248-335-35
11	56.4	8.0	748	1	US-08-361-467B-3
12	56.4	8.0	748	1	US-08-484-332C-3
13	56.4	8.0	5173	1	US-08-242-677-1
14	56	7.9	1364	1	US-08-265-087-3
15	56	7.9	1364	1	US-08-621-493-3
16	56	7.9	1364	2	US-08-965-688-3
17	55.8	7.9	1134	3	US-09-248-335-29
18	55.6	7.9	1172	1	US-07-945-288-9
19	55.6	7.9	1172	1	US-08-462-831-9
20	55.6	7.9	1172	1	US-08-461-809-9
21	55.6	7.9	1172	1	US-08-461-809-9
22	55.6	7.9	1172	4	PCT-US93-08518-9
23	55.4	7.9	1066	1	US-08-157-101A-4
24	55.4	7.9	1279	3	US-08-248-335-25
25	55.2	7.8	368	1	US-08-562-257-20
26	55.2	7.8	568	2	US-08-562-298-20
27	54.8	7.8	458	1	US-08-524-757-1
28	54.8	7.8	536	1	US-08-341-568-1

29	54.8	7.8	536	2	US-08-911-020-1	Sequence 1, Appli
30	54.8	7.8	688	5	5498694-3	Patent No. 5498694
31	54.8	7.8	976	2	US-08-504-459-9	Sequence 9, Appli
32	54.8	7.8	1534	1	US-08-300-903A-6	Sequence 6, Appli
33	54.6	7.7	222	3	US-08-481-190-15	Sequence 15, Appli
34	54.6	7.7	222	4	PCN-US93-00869-15	Sequence 5, Appli
35	54.4	7.7	140	1	US-08-628-417-5	Sequence 6, Appli
36	54.4	7.7	240	1	US-08-628-417-6	Sequence 24, Appli
37	54.4	7.7	1493	1	US-08-340-820-24	Sequence 24, Appli
38	54.4	7.7	1493	1	US-08-593-535-24	Sequence 3, Appli
39	54.2	7.7	289	1	US-08-341-568-3	Sequence 3, Appli
40	54.2	7.7	289	2	US-08-911-020-3	Sequence 1, Appli
41	54.2	7.7	1882	3	US-09-370-253-1	Sequence 25, Appli
42	54	7.7	991	3	US-08-924-747-25	Sequence 4, Appli
43	54	7.7	1721	4	PCT-US96-00994-3	Patent No. 5258287
44	54	7.7	2550	5	5258287-23	Sequence 4, Appli
45	53.8	7.6	1700	2	US-08-897-340-4	

#### ALIGNMENTS

RESULT 1  
US-08-909-965C-9  
Sequence 9, Application US/08909965C  
Patent No. 536078  
GENERAL INFORMATION:  
APPLICANT: Kuga Tetsuo  
APPLICANT: Nakagawa Satoshi  
APPLICANT: Sakaki Yoshiyuki  
APPLICANT: Zhao Nanding  
APPLICANT: Hashida Hideji  
TITLE OF INVENTION: NOVEL DNA, NOVEL POLYPEPTIDE  
TITLE OF INVENTION: AND NOVEL ANTIBODY  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FITZPATRICK, CELLA, HARPER AND SCINTO  
STREET: 277 Park Avenue  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10172-0194  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/909, 965C  
FILING DATE: August 12, 1997  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 322745/95  
APPLICATION NUMBER: PCT/JP96/03630  
FILING DATE: 12-Dec-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Lawrence S. Perry  
REGISTRATION NUMBER: 31865  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-758-2400  
TELEFAX: 212-758-2962  
TELEX: 236262  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 882 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
ORIGINAL SOURCE: human  
ORGANISM: human  
IMMEDIATE SOURCE:



QY	Db	QY	Db
611	1797	671	1857
CTCTCTCAAGACATCTCAGCGTATATATTAACCTGACGACCAAAAAA	TATTTCACATTTTATTAATTAATGGAAGCTACTATCATTTATATGCAAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAA
6/0	1856	705	1891

```

RESULT      6
US-08-545-196B-10
; Sequence 10, Application US/08545196B
; Patent No. 6080577
; GENERAL INFORMATION:
; APPLICANT: MELKI, JUDITH
; APPLICANT: MUNNICH, ARNOLD
; TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
; TITLE OF INVENTION: FOR SPINAL MUSCULAR ATROPHY
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/545,196B
; FILING DATE: 19-OCT-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FARACI, C. J.
; REGISTRATION NUMBER: 32,350
; REFERENCE/DOCKET NUMBER: 2121-110P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1582 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-545-196B-10

Query Match          8.2%; Score 57.6; DB 3; Length 1582;
Best Local Similarity 75.0%; Pred. No. 0.00014;
Matches 72; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 610 gtcttcaacaatatttaagtgtagcgtgatatctaactgaacacaaaaaa 669
    ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1461 GTATGTAAATAAATAATTAACTTTTTTYTAATAAAAAAAAAAAAAAAA 1520
QY 670 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 705
DB 1521 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1556

RESULT      7
US-08-545-196B-12
; Sequence 12, Application US/03545196B
; Patent No. 6080577
; GENERAL INFORMATION:
; APPLICANT: MELKI, JUDITH
; APPLICANT: MUNNICH, ARNOLD
; TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE
; TITLE OF INVENTION: FOR SPINAL MUSCULAR ATROPHY
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP
; STREET: PO BOX 747
; CITY: FALLS CHURCH
; STATE: VA
; COUNTRY: USA

```

```

1 ZIP: 22040-0747
2
3 COMPUTER READABLE FORM:
4
5 MEDIUM TYPE: Floppy disk
6
7 COMPUTER: IBM PC compatible
8
9 OPERATING SYSTEM: PC-DOS/MS-DOS
10
11 SOFTWARE: PatentIn Release #1.0, Version #1.30
12
13 CURRENT APPLICATION DATA:
14
15 APPLICATION NUMBER: US/08/545,196B
16
17 FILING DATE: 19-OCT-1995
18
19 CLASSIFICATION: 435
20
21 ATTORNEY/AGENT INFORMATION:
22
23 NAME: PARACI, C. J.
24
25 REGISTRATION NUMBER: 32,350
26
27 REFERENCE/DOCKET NUMBER: 2121-110P
28
29 TELECOMMUNICATION INFORMATION:
30
31 TELEPHONE: (703) 205-8000
32
33 TELEFAX: (703) 205-8050
34
35 INFORMATION FOR SEO ID NO: 12:
36
37 SEQUENCE CHARACTERISTICS:
38
39 LENGTH: 1582 base pairs
40
41 TYPE: nucleic acid
42
43 STRANDEDNESS: double
44
45 TOPOLOGY: linear
46
47 MOLECULE TYPE: CDNA
48
49 US-08-545-196B-12

```

[illegible]

```

RESULT      8
US-08-841-349-18/c
: Sequence 18, Application US/08041349B
: Patent No. 5955594
: GENERAL INFORMATION:
: APPLICANT: MISHRA, LOPA
: TITLE OF INVENTION: GENES CODING PROTEINS FOR EARLY LIVER DEVELOPMENT. . .
: FILE REFERENCE: XX/PO4470US0
: CURRENT APPLICATION NUMBER: US/08/841,349B
: CURRENT FILING DATE: 1997-04-30
: NUMBER OF SEQ ID NOS: 18
: SOFTWARE: Patentin Ver. 2.0
: SEQ ID NO 18
:   LENGTH: 467
:   TYPE: DNA
: ORGANISM: Mus musculus
: FEATURE:
: OTHER INFORMATION: For all n's in this sequence, n=(a or g or c or t)
US-08-841-349-18

```

	Query Match	8.1%	Score 57.4	DB 2	Length 467	
	Best Local Similarity	58.6%	Pred. No. 0.0001			
	Matches	85	Conservative	0	Mismatches 60	Indels 0
Qy	556	ggccctcagcggccccgcagcgtgcctcttcccaacccctctgtaagtaacccctgttctcta	617			
Db	152	GGGNCAGGCGCTGTGGCTTCCTTCTTAAACCCTCCGCCCTTGGAANNNTTNTATAAAAA	93			
Qy	618	aacactatttaagtgaigtatattataaactgtgacacaaaaaanaaaaaaa	677			
Db	92	AAAAAACTTTAAANNNTTNTNNNNNNAAAAAANAANAANAANAANAANAANAANAANAANA	33			





FILING DATE: 17-JUN-1994  
 APPLICATION NUMBER: US 08/229,282  
 FILING DATE: 18-APR-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Bak, Mary E.  
 REGISTRATION NUMBER: 31,215  
 REFERENCE/DOCKET NUMBER: WST51AUSA  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 215-540-9206  
 TELEFAX: 215-540-5818  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1364 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: unknown  
 MOLECULE TYPE: CDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 101..859  
 US-08-621-493-3

Query Match	7.9%	Score 56:	DB 1:	Length 1364:
Best Local Similarity	74.0%	Pred. No.	0.0003:	
Matches 71:	Conservative	0:	Mismatches 25:	Indels 0:
				Gaps 0:

Qy	Db
610	610
gttccttaaacacttcctttaaagtgatcgcggtcttcttaacgcgtgaacccaaaaaa	669
1267	1267
GTGATTTTGAATAATTTTAAAGCTCTCTAAATTAATAAGATTTGATTTAAAAAAA	1326
670	705
aaaaaaaaaaaaaaaaaaaaaaaaaaaaa	
1327	1362
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	







PF 27-MAY-1999; 99WO-US11644.  
 XX  
 PR 29-MAY-1998; 98US-0087340.  
 PR 10-SEP-1998; 98US-0099805.  
 PR 30-APR-1999; 99US-0131965.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Ruben SM, Edner R;  
 XX WPI: 2000-072622/06.  
 DR N-PSDB: Z36834.  
 XX  
 PT Novel polynucleotides used to develop products for treating e.g. immune  
 PT disorders, blood disorders, autoimmune disorders, allergies,  
 PT inflammation, hyperproliferative disorders or infections  
 XX  
 PS Claim 25; Fig 1; 170pp; English.  
 XX  
 CC The present sequence represents a partial human interleukin-21 (IL-21)  
 CC protein. The specification also describes IL-22 polynucleotides and  
 CC polypeptides. The IL-21 polynucleotide was isolated from a cDNA library  
 CC of apoptotic T-cells. IL-21 and IL-22 may be useful in treating  
 CC deficiencies or disorders of the immune system, by activating or  
 CC inhibiting the proliferation, differentiation, or mobilization  
 CC (chemotaxis) of immune cells, treating or detecting deficiencies or  
 CC disorders of haematopoietic cells, to modulate haemostatic or  
 CC thrombolytic activity, in treating or detecting autoimmune disorders,  
 CC treating asthma (particularly allergic asthma) or other respiratory  
 CC problems, to treat and/or prevent organ rejection or graft-versus-host  
 CC disease (GVHD), to modulate inflammation, to treat or detect  
 CC hyperproliferative disorders, to treat or detect infectious agents, to  
 CC of tissues, IL-21 and IL-22 may also increase or decrease the  
 CC differentiation or proliferation of embryonic stem cells and  
 CC haematopoietic lineage, may be used to modulate mammalian  
 CC characteristics.  
 CC  
 XX Sequence 87 AA;  
 SQ

Query Match 100.0%; Score 463; DB 21; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-52;  
 Matches 87; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ARVDDEDYRYPQKLAFAECLRGCIDARTGRTAALNSVRLQSLVLRRRPCSRDGSGL 60  
 DB 1 arvdededypqklafaeclrgcidartgretaalnsrvllqslvlrrpcsrdsqsl 60  
 QY 61 PTPGAFARFTEFIHVPVGGCTGVLPNSV 87  
 DB 61 ptpgafarftefihvpvgctgvlpnsv 87

RESULT 2  
 Y92238  
 ID Y92238 standard; Protein; 197 AA.  
 XX  
 AC Y92238;  
 XX  
 DT 10-AUG-2000 (first entry)  
 XX  
 DE Human interleukin-17 (IL-17) homologue.  
 XX  
 KW Interleukin 17; IL-17; haematopoiesis; chemotherapy; cytostatic;  
 KW antianemic; cardiant; hemostatic; anti-inflammatory; anti-HIV.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FH Peptide 1..18  
 FT /label= signal\_peptide  
 FT Protein 19..197

FT /label= mature\_protein  
 XX WO200020593-A1.  
 PN 13-APR-2000.  
 XX  
 PD 30-SEP-1999; 99WO-US22678.  
 XX  
 PF 02-OCT-1998; 98US-0102883.  
 PR 01-DEC-1998; 98US-0110405.  
 PR 11-JUN-1999; 99US-0138910.  
 XX  
 PA (ELIL) LILLY & CO ELI.  
 XX  
 PI Glasebrook AL, Su EW, Wei J, Liu L;  
 XX WPI: 2000-303778/26.  
 DR N-PSDB: A09153.  
 DR  
 DR  
 XX  
 PS Claim 16; Page 92-93; 11pp; English.  
 XX  
 CC Interleukin 17 (IL-17) stimulates hematopoiesis and production of  
 CC neutrophils, granulocytes, or platelets, this may be useful during  
 CC chemotherapy. IL-17 homologues have at least one actively selected  
 CC from induction of cytotoxic T cells. Induction of lymphokine-activated  
 CC killer cell proliferation or a B or T cell stimulation. The IL-17  
 CC homologue may also be used to treat viral or bacterial infections,  
 CC immune related diseases, anemia, leukemia, thrombocytopenia, uremia,  
 CC Von Willebrand disease, postoperative cardiovascular dysfunction,  
 CC treatment of AIDS (acquired immune deficiency syndrome)-related bone  
 CC marrow failure, and inflammatory diseases of the gastrointestinal  
 CC system, joints, and lungs.  
 CC  
 XX Sequence 197 AA;  
 SQ

Query Match 99.1%; Score 459; DB 21; Length 197;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-51;  
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RVDDEDYRYPQKLAFAECLRGCIDARTGRTAALNSVRLQSLVLRRRPCSRDGSGLP 61  
 DB 112 rvdededypqklafaeclrgcidartgretaalnsrvllqslvlrrpcsrdsqslp 171  
 QY 62 TPGAFARFTEFIHVPVGGCTGVLPNSV 87  
 DB 172 tpgafarftefihvpvgctgvlpnsv 197

RESULT 3  
 Y44460  
 ID Y44460 standard; Protein; 197 AA.  
 XX  
 AC Y44460;  
 XX  
 DT 27-MAR-2000 (first entry)  
 XX  
 DE Human Interleukin 17C, PRO1122 polypeptide.  
 XX  
 KW Interleukin; IL-17C; PRO1122 polypeptide; clone DNA62377-1381-1; UNO561;  
 KW cytokine IL-17; cytotoxic T-lymphocyte-associated antigen 8; CTLA-8;  
 KW hybridisation probe; antagonist; degenerative cartilaginous disorder;  
 KW agonist; diagnosis; therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FH Peptide 1..18  
 FT /label= signal\_peptide  
 FT Protein 19..197



Query Match	99.1%: Score 459; DB 21; Length 197;
Best Local Similarity	100.0%; Pred. No. 3.3e-51;
Matches	86; Conservative 0; Mismatches 0; Indels 0; Gaps
Qy 2	RVDTEDEKRPKLFKFAECLRCGICIDARTGRTAALNSVRLQSLDLVRRRPGSRDGSGLP 61
Db 112	rvtddedtyppkklfaeclrcgcidartgretaaInsvrlqlslvllrrrrpcsrdsqslp 171
Qy 62	TPGAFATFTEFLHPVVGSTCYLPRSV 87
Db 172	lpgatafhteifhvpvgctcvlprsv 197
RESULT 5	
ID Y44485	Y44485 standard; Protein; 206 AA.
AC Y44485;	
DT 27-MAR-2000	(first entry)
XX Human Interleukin 17C with C-terminal Gly(His)8 tag, IL-17C.his.	
DE Interleukin; IL-17C.his; PRO1122 polypeptide; clone DMA62377-1381-1;	
KW immunoprecipitation; IL-17 receptor extracellular domain; IL-17R ECD;	
RV cytokine IL-17; hybridisation probe; antagonist; Gly(His)8 tag; agonist;	
XX degenerative cartilaginous disorder; diagnosis; therapy.	
XX Homo sapiens.	
XX Key	Location/Qualifiers
FT Peptide	1..18
FT /label=	Signal_peptide
FT Protein	19..197
FT /label=	Mature_IL-17C_polypeptide
FT /note=	"used to treat degenerative cartilaginous disorder"
FT Misc-difference 109	
FT /note=	"Conserved Trp residue"
FT Misc-difference 129	
FT /note=	"Conserved Cys residue"
FT Misc-difference 134	
FT /note=	"Conserved Cys residue"
FT Misc-difference 163	
FT /note=	"Conserved Cys residue"
FT Misc-difference 189	
FT /note=	"Conserved Cys residue"
FT Misc-difference 191	
FT /note=	"Conserved Cys residue"
FT Misc-difference 198..206	
FT /note=	"C-terminal Gly(His)8 tag"
XX W09960127-A2.	
XX 25-NOV-1999.	
XX 14-MAY-1999;	99WO-US10733.
XX 15-MAY-1998;	98US-0085579.
XX 23-DEC-1998;	98US-0113621.
XX (GETH ) GENENTECH INC.	
XX Chen J, Filvarcoff E, Goddard A, Gurney AL, Li H, Wood WI;	
XX WPI; 2000-116314/10.	
XX New polypeptides designated PRO1031 and PRO1122 used to treat a	
XX degenerative cartilaginous disorder -	
XX Example 11; Page 138-139; 141pp; English.	

XX	The present sequence is the human PRO1122 polypeptide, with a C-terminal			
CC	Gly(His)8 tag, IL-17C.f.c; derived from the clone DNA62377-1381-1.			
CC	This sequence is used in a competitive binding experiment for the			
CC	immunoprecipitation of IL-17 receptor extracellular domain (IL-17R ECD).			
CC	The entire coding region of IL-17C can be used as hybridisation probe.			
CC	The PRO1122 polypeptide, agonist or antagonist, is used to diagnose and			
CC	treat a degenerative cartilaginous disorder.			
XX	Sequence	206 AA:		
SO				
Query Match	99.1%;	Score 459;	DB 21; Length 206;	
Best Local Similarity	100.0%;	Pred. No. 3.5e-51;		
Matches	86; Conservative	0; Mismatches	0; Indels 0; Gaps	
OY	2 RVDDDEDHYQKIAIPFECICRGCIDARCRGRFMAINSVRLIQSLTLVRRPSCRDGSGIP 61			
Db	112 rvdcdetrpqklatfaecclercgldarfyretaaInsvrllqslvltlrpcsdgsgip 171			
OY	62 TPQAFARHTETFIHPVGTCTCVPRSY 87			
Db	172 tpgafafhteflnvpygccvprsy 197			
RESULT	6			
ID	Y44462			
XX	Y44462 standard; Protein: 425 AA.			
XX	Y44462;			
XX	AC			
XX	27-MAR-2000 (first entry)			
DT				
DE	Human Interleukin 17C-IgG1 Fc fusion protein, hIL-17C.f.c.			
XX				
KW	Interleukin; IL-17C.f.c; fusion protein; PRO1122 polypeptide; cytokine;			
KW	human IgG1; fluorescence-activated cell sorter analysis; FACS;			
KW	Tumour Necrosis Factor-alpha; TNF-alpha; leukemic monocyte; THP-1 cell.			
XX				
OS	Homo sapiens.			
XX				
FH	Key	Location/Qualifiers		
FT	Peptide	1..18		
FT		/label= Signal_peptide		
FT	Protein	19..197		
FT		/label= Mature_IL-17C-polypeptide		
FT		/note= "used to treat degenerative cartilaginous disorder"		
FT	Misc-difference	109		
FT		/note= "Conserved Trp residue"		
FT	Misc-difference	129		
FT		/note= "Conserved Cys residue"		
FT	Misc-difference	134		
FT		/note= "Conserved Cys residue"		
FT	Misc-difference	163		
FT		/note= "Conserved Cys residue"		
FT	Misc-difference	189		
FT		/note= "Conserved Cys residue"		
FT	Misc-difference	191		
FT		/note= "Conserved Cys residue"		
FT	Region	197..425		
FT		/note= "Sequence derived from Fc region of human IgG1"		
XX				
PN	W09960127-A2.			
XX				
PD	25-NOV-1999.			
XX				
PF	14-MAY-1999;	99WO-US10733.		
XX				
PR	15-MAY-1998;	98US-0085579.		
XX				
RR	23-DEC-1998;	98US-0113621.		
XX				
PA	(GETH ) GENENTECH INC.			

XX Chen J, Filvaroff E, Goddard A, Gurney AL, Li H, Wood WT;  
 XX WPI; 2000-116314/10.  
 DR  
 XX New polypeptides designated PRO1031 and PRO1122 used to treat a  
 PT degenerative cartilaginous disorder -  
 XX  
 PS Example 12: Page 129-130; 141pp; English.  
 XX  
 CC The present sequence is the human IL-17C.fc fusion protein, derived from  
 CC PRO1122 polypeptide and the Fc region of human IgG1. The cytokine IL-17C  
 CC can be used to induce the release of TNF-alpha from human leukemic  
 CC monocytic, THP-1 cells. The fusion protein, IL-17C.fc is used to identify  
 CC the binding of IL-17C to THP-1 cells, using fluorescence-activated cell  
 CC sorter analysis (FACS).  
 XX  
 SQ Sequence 425 AA:  
 XX  
 QY Query Match 99.1%; Score 459; DB 21; Length 425;  
 Best Local Similarity 100.0%; Pred. No. 8.1e-51;  
 Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 DB 2 RVTDEDRYPRQKLAFAECICRCGICDARTGRETALNSVRLQSLVLRPRPCSRDGSGLP 61  
 112 rvtdeedrypqkrlataecrcgcidartgretalnsvrlqslvlrrpcsrdsqslp 171  
 QY 62 TPGAFAFHTEFIHVPVGCVCVLRPSV 87  
 172 tpgafafhfeihvpgcvcvlprsv 197  
 DB  
 RESULT 7  
 Y53891  
 ID Y53891 standard; Protein; 160 AA.  
 XX  
 AC Y53891;  
 XX  
 DT 13-MAR-2000 (first entry)  
 XX  
 DE Partial amino acid sequence of human interleukin-22.  
 XX  
 KW Human; interleukin-22; IL-22; IL-21; immune system disorder;  
 KW immune cell chemotaxis; haematopoietic cell disorder;  
 KW haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;  
 KW respiratory problem; organ rejection; graft-versus-host disease; GVHD;  
 KW inflammation; hyperproliferative disorder; tissue regeneration;  
 KW embryonic stem cell differentiation; embryonic stem cell proliferation;  
 KW haematopoietic lineage; allergic asthma.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 26 /note= "potential N-linked glycosylation site"  
 FT Domain 57..64 /note= "conserved domain I"  
 FT Domain 72..77 /note= "conserved domain II"  
 FT Domain 99..105 /note= "conserved domain III"  
 FT Domain 121..128 /note= "conserved domain IV"  
 FT Modified-site 139 /note= "potential N-linked glycosylation site"  
 FT  
 FT  
 FT  
 PN WO9961617-A1.  
 XX  
 PD 02-DEC-1999.  
 XX  
 PF 27-MAY-1999; 99WO-US11644.  
 XX

PR 29-MAY-1998; 98US-0087240.  
 PR 10-SEP-1998; 98US-0099805.  
 PR 30-APR-1999; 99US-0131965.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Ruben SM, Ebner R;  
 XX  
 DR WPI; 2000-072622/06.  
 DR N-PSDB; Z36835.  
 XX  
 PT Novel polynucleotides used to develop products for treating e.g. immune  
 PT disorders, blood disorders, autoimmune disorders, allergies,  
 PT inflammation, hyperproliferative disorders or infections -  
 XX  
 PS Claim 27; Fig 2A-B; 170pp; English.  
 XX  
 CC The present sequence represents a partial human interleukin-22 (IL-22)  
 CC protein. The specification also describes IL-21 polynucleotides and  
 CC polypeptides. The IL-21 polynucleotide was isolated from a cDNA library  
 CC of epileptic frontal cortex. IL-21 and IL-22 may be useful in treating  
 CC deficiencies or disorders of the immune system, by activating or  
 CC inhibiting the proliferation, differentiation, or mobilization  
 CC (chemotaxis) of immune cells, treating or detecting deficiencies or  
 CC disorders of haematopoietic cells, to modulate haemostatic or  
 CC thrombolytic activity, in treating or detecting autoimmune disorders,  
 CC treating asthma (particularly allergic asthma) or other respiratory  
 CC problems, to treat and/or prevent organ rejection or graft-versus-host  
 CC disease (GVHD), to modulate inflammation, to treat or detect  
 CC hyperproliferative disorders, to treat or detect infectious agents, to  
 CC differentiate, proliferate and attract cells, leading to the  
 CC regeneration of tissues, IL-21 and IL-22 may also increase or decrease  
 CC the differentiation or proliferation of embryonic stem cells and  
 CC haematopoietic lineage, may be used to modulate mammalian  
 CC characteristics.  
 XX  
 SQ Sequence 160 AA:  
 XX  
 QY Query Match 31.0%; Score 143.5; DB 21; Length 160;  
 Best Local Similarity 37.8%; Pred. No. 7.7e-11;  
 Matches 31; Conservative 10; Mismatches 32; Indels 9; Gaps 2;  
 DB 2 RVTDEDRYPRQKLAFAECICRCGICDARTGRETALNSVRLQSLVLRPRPCSRDGSGLP 61  
 55 rlsydparyprylpeeyclorgcltqltgsedvfrsapymplylrrtpacagrsv- 113  
 QY 62 TPGAFAFHTE-FIHVPVGCVCV 82  
 114 -----yteaayvlipgctcv 128  
 DB  
 RESULT 8  
 Y53893  
 ID Y53893 standard; Protein; 173 AA.  
 XX  
 AC Y53893;  
 XX  
 DT 13-MAR-2000 (first entry)  
 XX  
 DE Partial amino acid sequence of human interleukin-22.  
 XX  
 KW Human; interleukin-22; IL-22; IL-21; immune system disorder;  
 KW immune cell chemotaxis; haematopoietic cell disorder;  
 KW haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;  
 KW respiratory problem; organ rejection; graft-versus-host disease; GVHD;  
 KW inflammation; hyperproliferative disorder; tissue regeneration;  
 KW embryonic stem cell differentiation; embryonic stem cell proliferation;  
 KW haematopoietic lineage; allergic asthma.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers



XX AC Y70655;  
XX 18-JUL-2000 (first entry)  
XX DE Mature human transforming growth factor beta-9, ztgbeta-9 protein-2.  
XX DE Human transforming growth factor beta-9; ztgbeta-9;  
XX KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
XX KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
XX KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
XX KW antiviral; cyostatic.  
XX OS Homo sapiens.  
XX PN WO200015798-A2.  
XX PD 23-MAR-2000.  
XX PF 17-SEP-1999; 99WO-US21677.  
XX PR 17-SEP-1998; 98US-0154817.  
XX PA (ZYMO ) ZYMOGENETICS INC.  
XX PI Presnell SR, Taft DW, Foley KP;  
XX DR WPI: 2000-271436/23.  
XX DR N-PSDB: Z52195.  
XX PT Polynucleotides encoding a novel transforming growth factor beta-9  
XX PT polypeptide, designated ztgb beta-9, useful as an antiviral and  
XX PT antiproliferative agent -  
XX PS Claim 6; Page 84; 97pp; English.  
XX PS The present sequence is the mature human transforming growth factor  
XX CC beta-9, designated ztgb beta-9. This is a mature sequence excluding the  
XX CC signal sequence extending from amino acid 17 to and including amino acid  
XX CC 202 of ztgb beta-9. Human ztgb beta-9 was isolated from an arrayed  
XX CC pituitary gland cDNA plasmid library by PCR screening. This can be used  
XX CC to treat a variety of neurodegenerative diseases such as amyotrophic  
XX CC lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,  
XX CC Parkinson's disease and peripheral neuropathies, or demyelinating  
XX CC diseases including multiple sclerosis. Ztgb beta-9 peptides have  
XX CC antiviral activity and may also be used to regulate the proliferation,  
XX CC differentiation and apoptosis of neurons, glial cells, lymphocytes,  
XX CC hematopoietic cells and stromal cells.  
XX SQ Sequence 186 AA;  
SQ  
Query Match 31.0%; Score 143.5; DB 21; Length 186;  
Best Local Similarity 37.8%; Pred. No. 9, 1e-11;  
Matches 31; Conservative 10; Mismatches 32; Indels 9; Gaps 2;  
QY 2 RVDTDDEDRYPKLIARFACRCGICDARTGRTAALNSVRLLOSILVLRPPCSRDSSGLP 61  
DB 81 risydparyprylpeayclrcgcltglfgedvrtfsapvymptvlrrtpacagrsv- 139  
QY 62 TPGAFAFHTE-FRIHPVGCTCV 82  
DB 140 -----yteayvtilpvgtcv 154  
RESULT 11  
ID Y70654  
AC Y70654 standard; Protein: 187 AA.  
XX Y70654;  
XX 18-JUL-2000 (first entry)  
XX

DE DE Mature human transforming growth factor beta-9, ztgbeta-9 protein-1.  
XX KW Human transforming growth factor beta-9; ztgbeta-9;  
XX KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
XX KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
XX KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
XX KW antiviral; cyostatic.  
XX OS Homo sapiens.  
XX PN WO200015798-A2.  
XX PD 23-MAR-2000.  
XX PF 17-SEP-1999; 99WO-US21677.  
XX PR 17-SEP-1998; 98US-0154817.  
XX PA (ZYMO ) ZYMOGENETICS INC.  
XX PI Presnell SR, Taft DW, Foley KP;  
XX DR WPI: 2000-271436/23.  
XX DR N-PSDB: Z52195.  
XX PT Polynucleotides encoding a novel transforming growth factor beta-9  
XX PT polypeptide, designated ztgb beta-9, useful as an antiviral and  
XX PT antiproliferative agent -  
XX PS Claim 6; Page 84; 97pp; English.  
XX PS The present sequence is the mature human transforming growth factor  
XX CC beta-9, designated ztgb beta-9. This is a mature sequence excluding the  
XX CC signal sequence extending from amino acid 16 to and including amino acid  
XX CC 202 of ztgb beta-9. Human ztgb beta-9 was isolated from an arrayed  
XX CC pituitary gland cDNA plasmid library by PCR screening. This can be used  
XX CC to treat a variety of neurodegenerative diseases such as amyotrophic  
XX CC lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,  
XX CC Parkinson's disease and peripheral neuropathies, or demyelinating  
XX CC diseases including multiple sclerosis. Ztgb beta-9 peptides have  
XX CC antiviral activity and may also be used to regulate the proliferation,  
XX CC differentiation and apoptosis of neurons, glial cells, lymphocytes,  
XX CC hematopoietic cells and stromal cells.  
XX SQ Sequence 187 AA;  
SQ  
Query Match 31.0%; Score 143.5; DB 21; Length 187;  
Best Local Similarity 37.8%; Pred. No. 9, 2e-11;  
Matches 31; Conservative 10; Mismatches 32; Indels 9; Gaps 2;  
QY 2 RVDTDDEDRYPKLIARFACRCGICDARTGRTAALNSVRLLOSILVLRPPCSRDSSGLP 61  
DB 82 risydparyprylpeayclrcgcltglfgedvrtfsapvymptvlrrtpacagrsv- 140  
QY 62 TPGAFAFHTE-FRIHPVGCTCV 82  
DB 141 -----yteayvtilpvgtcv 155  
RESULT 12  
ID Y70663  
AC Y70663 standard; Protein: 187 AA.  
XX Y70663;  
XX 18-JUL-2000 (first entry)  
XX DE Mature human ztgbeta-9 variant protein.  
XX DE Human transforming growth factor beta-9; ztgb beta-9;  
XX KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
XX KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
XX

KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
KM antiviral; cytosolic.  
XX  
OS Homo sapiens.  
XX  
PN W0200015798-A2.  
XX  
PD 23-MAR-2000.  
XX  
PE 17-SEP-1999; 99WO-US21677.  
XX  
PR 17-SEP-1998; 98US-0154817.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Presnell SR, Taft DW, Foley KP;  
XX  
DR WPI: 2000-271436/23.  
DR N-PSDB: 252201.  
XX  
PT Polynucleotides encoding a novel transforming growth factor beta-9  
PT polypeptide, designated ztgf beta-9, useful as an antiviral and  
PT antiproliferative agent -  
XX  
PS Claim 6; Page 93-94; 97pp; English.  
XX  
CC The present sequence is a mature variant human transforming growth  
CC factor beta-9 protein, designated ztgf beta-9. Human ztgf beta-9 was  
CC isolated from an arrayed pituitary gland cDNA plasmid library by PCR  
CC screening. This can be used to treat a variety of neurodegenerative  
CC diseases such as amyotrophic lateral sclerosis (ALS), Alzheimer's  
CC disease, Huntington's disease, Parkinson's disease and peripheral  
CC neuropathies, or demyelinating diseases including multiple sclerosis.  
CC ztgf beta-9 peptides have antiviral activity and may also be used to  
CC regulate the proliferation, differentiation and apoptosis of neurons,  
CC glial cells, lymphocytes, hematopoietic cells and stromal cells.  
XX  
SQ Sequence 187 AA;

Query Match 31.0%; Score 143.5; DB 21; Length 187;  
Best Local Similarity 37.8%; Pred. No. 9.2e-11;  
Matches 31; Conservative 10; Mismatches 32; Indels 9; Gaps 2;

QY 2 RVDDDEDYRYPQKLAFAECIDRGCIIDARTGRTAALNSVRLQSLVLRPRPCSRDGSGLP 61  
| : | ||| : | | ||||| : | : : |||| | :  
Db 82 risypdarpyprrlpeaylclcgcltglfgeedvfrfsapymptvlrrtpacaggrsv- 140  
: | : : |||||  
QY 62 TPGAFAFHTF-FIHVPVGCCTCV 82  
: | : : |||||  
Db 141 -----yleayvtlipygctcv 155

RESULT 13  
Y70653  
ID Y70653 standard; Protein; 202 AA.  
XX  
AC Y70653;  
XX  
DT 18-JUL-2000 (first entry)  
XX  
DE Human transforming growth factor beta-9, ztgf beta-9 protein.  
XX  
KW Human transforming growth factor beta-9; ztgf beta-9;  
KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
KW antiviral; cytosolic.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..15

FT  
FT Protein /label= signal\_peptide  
FT 16..202  
/note= "Mature transforming growth factor beta-9"  
XX  
XX  
PN W0200015798-A2.  
XX  
PD 23-MAR-2000.  
XX  
PE 17-SEP-1999; 99WO-US21677.  
XX  
PR 17-SEP-1998; 98US-0154817.  
XX  
PA (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Presnell SR, Taft DW, Foley KP;  
XX  
DR WPI: 2000-271436/23.  
DR N-PSDB: 252195.  
XX  
PT Polynucleotides encoding a novel transforming growth factor beta-9  
PT polypeptide, designated ztgf beta-9, useful as an antiviral and  
PT antiproliferative agent -  
XX  
PS Claim 6; Page 83; 97pp; English.  
XX  
CC The present sequence is the human transforming growth factor beta-9,  
CC designated ztgf beta-9. Human ztgf beta-9 was isolated from an arrayed  
CC pituitary gland cDNA plasmid library by PCR screening. This can be used  
CC to treat a variety of neurodegenerative diseases such as amyotrophic  
CC lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,  
CC Parkinson's disease and peripheral neuropathies, or demyelinating  
CC diseases including multiple sclerosis. ztgf beta-9 peptides have  
CC antiviral activity and may also be used to regulate the proliferation,  
CC differentiation and apoptosis of neurons, glial cells, lymphocytes,  
CC hematopoietic cells and stromal cells.  
XX  
SQ Sequence 202 AA;

Query Match 31.0%; Score 143.5; DB 21; Length 202;  
Best Local Similarity 37.8%; Pred. No. 1e-10;  
Matches 31; Conservative 10; Mismatches 32; Indels 9; Gaps 2;

QY 2 RVDDDEDYRYPQKLAFAECIDRGCIIDARTGRTAALNSVRLQSLVLRPRPCSRDGSGLP 61  
| : | ||| : | | ||||| : | : : |||| | :  
Db 97 risypdarpyprrlpeaylclcgcltglfgeedvfrfsapymptvlrrtpacaggrsv- 155  
: | : : |||||  
QY 62 TPGAFAFHTF-FIHVPVGCCTCV 82  
: | : : |||||  
Db 156 -----yleayvtlipygctcv 170

RESULT 14  
Y70662  
ID Y70662 standard; Protein; 209 AA.  
XX  
AC Y70662;  
XX  
DT 18-JUL-2000 (first entry)  
XX  
DE Human transforming growth factor beta-9, ztgf beta-9 variant protein.  
XX  
KW Human transforming growth factor beta-9; ztgf beta-9;  
KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
KW antiviral; cytosolic.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..22  
/label= signal\_peptide







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2001, 12:49:25 ; Search time 38.3 Seconds  
(without alignments)  
266.243 Million cell updates/sec

Title: US-09-320-713-2

Perfect score: 463

Sequence: 1 ARVDDEDHYPOKLAFAECL.....FHTEFIHVPGCTVLPNSV 87

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL.15:\*  
2: sp.archaea:\*  
3: sp.bacteria:\*  
4: sp.fungi:\*  
5: sp.human:\*  
6: sp.invertebrate:\*  
7: sp.mammal:\*  
8: sp.mhcc:\*  
9: sp.organelle:\*  
10: sp.phage:\*  
11: sp.plant:\*  
12: sp\_rodent:\*  
13: sp.virus:\*  
14: sp.vertibrate:\*  
15: sp.unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	459	99.1	197	4	Q9P0M4
2	136	29.4	78	4	Q9NUE6
3	123	26.6	151	12	O40633
4	110.5	23.9	180	11	Q9QXT6
5	98.5	21.3	180	4	Q9UHP5
6	88	19.0	148	5	Q19778
7	69.5	15.0	39	6	Q9TUP5
8	69.5	15.0	39	6	Q9TUP4
9	69.5	15.0	39	6	Q9TUP3
10	67	14.5	146	2	Q9S2C5
11	65	14.0	518	13	O13035
12	63.5	13.7	563	10	O24596
13	62.5	13.5	745	5	O9U195
14	62	13.4	1737	4	O75097
15	61	13.2	874	11	O9QTP0
16	61	13.2	2221	5	Q9U1M1
17	60.5	13.1	315	5	O9VUV7
18	60.5	13.1	227	4	O15935
19	60	13.0	227	4	O15935

20	60	13.0	245	5	O45998	O4598 caenorhabdi
21	60	13.0	436	2	O9XA68	O9XA68 streptomyc
22	60	13.0	812	5	Q9VT12	Q9VT12 drosophila
23	59.5	12.9	270	13	O93391	O93391 colurnix co
24	59.5	12.9	491	5	O18351	O18351 caenorhabdi
25	59.5	12.9	544	5	O20780	O20780 caenorhabdi
26	59.5	12.9	581	10	O65478	O65478 arbidopsi
27	59.5	12.9	743	12	O83111	O83111 mouse adeno
28	59.5	12.9	938	10	O9LDD7	O9LDD7 lycopersico
29	59	12.7	198	11	Q9JUE6	Q9JUE6 mus musculu
30	59	12.7	230	4	Q9NSC9	Q9NSC9 homo sapien
31	59	12.7	955	4	O99466	O99466 homo sapien
32	59	12.7	1999	4	O99940	O99940 homo sapien
33	59	12.7	2003	4	O00306	O00306 homo sapien
34	58.5	12.6	93	2	O9RIS4	O9RIS4 streptomyc
35	58.5	12.6	102	2	O9X8Y1	O9X8Y1 porcine ade
36	58.5	12.6	102	2	O9L1X5	O9L1X5 streptomyc
37	58.5	12.6	838	12	O41012	O41012 porcine ade
38	58.5	12.6	838	12	O9YTR7	O9YTR7 porcine ade
39	58.5	12.6	907	10	O64828	O64828 arbidopsi
40	58.5	12.6	2473	11	O9QZ84	O9QZ84 mus musculu
41	58	12.5	183	2	O9RZ51	O9RZ51 delinococcus
42	58	12.5	194	10	O22272	O22272 arbidopsi
43	58	12.5	347	4	O9Y6U7	O9Y6U7 homo sapien
44	58	12.5	524	2	O9RYR8	O9RYR8 delinococcus
45	58	12.5	709	12	O9JRE5	O9JRE5 trichomonas

## ALIGNMENTS

RESULT 1	Q9P0M4	PRELIMINARY;	PRT;	197 AA.
AC	Q9P0M4			
DT	01-OCT-2000 (TREMblrel. 15, Created)			
DT	01-OCT-2000 (TREMblrel. 15, Last sequence update)			
DT	01-OCT-2000 (TREMblrel. 15, Last annotation update)			
DE	INTERLEUKIN 17C.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Li H., Chen J., Huang A., Stinson J., Heldens S., Foster J., Dowd P.,			
RA	Gurney A.L., Wood W.I.;			
RT	"Cloning and Characterization of IL-17B and IL17C, Two New Members of			
RT	the IL-17 Cytokine Family";			
RL	Proc. Natl. Acad. Sci. U.S.A. 97:773-778(2000).			
DR	EMBL; AF152099; AAF28105.1; -			
SQ	SEQUENCE 197 AA; 21765 MW; BAE0152E18DE7D08 CRC64;			
Query Match	99.1%; Score 459; DB 4; Length 197;			
Best local similarity	100.0%; Pred. NO. 3.9e-49;			
Matches	86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	2 RVDDDEDHYPOKLAFAECLRCGIDARTGRTAALNSVRLQSLVLRPPCRSDGSLP 61			
DB	112 RVDDDEDHYPOKLAFAECLRCGIDARTGRTAALNSVRLQSLVLRPPCRSDGSLP 171			
QY	62 TPGAFAFTEFIHVPVGCCTVLPNSV 87			
DB	172 TPGAFAFTEFIHVPVGCCTVLPNSV 197			
RESULT 2	Q9NUE6	PRELIMINARY;	PRT;	78 AA.
ID	Q9NUE6			
AC	Q9NUE6			
DT	01-OCT-2000 (TREMblrel. 15, Created)			
DT	01-OCT-2000 (TREMblrel. 15, Last sequence update)			



Matches 26; Conservative 12; Mismatches 37; Indels 5; Gaps 2;

OY 3 VDTDDRYPOKLAFAECRCGICDARTGRTAALNSVRLQSLVLRPPCSRDGSGLT 62  
 DB 105 INHDSRIPVLPKRCCLCTGCVNFTMQEDRSMSVVF-SQVYRRRLCPPP-----PR 159

OY 63 PGAFAPHTFEIHPVGCCTCV 82  
 DB 160 TGPCRORAVMEIIVGCTCI 179

RESULT 6  
 O19778 PRELIMINARY; PRT; 148 AA.

AC O19778:  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JAN-1999 (TREMBLrel. 09, Last annotation update)  
 DE F25D1.3 PROTEIN.  
 GN F25D1.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94150718; PubMed=7906398;  
 RA Wilson R., Alnscough R., Anderson K., Baynes C., Berks M.,  
 RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,  
 RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,  
 RA Girden A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,  
 RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,  
 RA Lightning J., Lloyd C., Mcmuray A., Mortimore B., O'Callaghan M.,  
 RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,  
 RA Smalton N., Smith A., Sonhammer E., Staden R., Sulston J.,  
 RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,  
 RA Watson A., Weinstock L., Wilkinson-Sprat J., Wooldman P.,  
 RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.  
 elegans."  
 RT Nature 368:32-38(1994).  
 RL EMBL: Z73973; CAA98268.1; -;  
 DR EMBL: Z73973; CAA98268.1; -;  
 SQ SEQUENCE 148 AA; 16738 MW; ADB51F438DFB3940 CRC64;

Query Match 19.0%; Score 88; DB 5; Length 148;  
 Best Local Similarity 32.6%; Pred. NO. 0.0025;  
 Matches 28; Conservative 15; Mismatches 31; Indels 12; Gaps 4;

OY 1 ARVTDDEDRYPOKLAFAECRCGICDARTGRTAALNSVRLQSLVLRPPCSRDGSGLT 60  
 DB 59 SRVNOESREPLIAESVCLC-----RKSNGSGAFCMPIYRKVILIRVSCDR-STGL 111

OY 61 PTPGAFAPHTFEIHPVGCCTCVLPRS 86  
 DB 112 ---WNYVRSTELI--TVGCHSVLPRT 132

RESULT 7  
 O9TUP5 PRELIMINARY; PRT; 39 AA.

AC O9TUP5:  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE INTERLEUKIN 17 (FRAGMENT).  
 GN IL-17.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Laurent F., Lacroix S.M., Mancassola R., Naciri M.;  
 RT "Partial sequence of bovine IL-17."  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF166487; AAD46376.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 39 AA; 4394 MW; 4A390A22973482A3 CRC64;

Query Match 15.0%; Score 69.5; DB 6; Length 39;  
 Best Local Similarity 42.5%; Pred. NO. 0.13;  
 Matches 17; Conservative 7; Mismatches 15; Indels 1; Gaps 1;

OY 9 RYPOKLAFAECRCGICDARTGRTAALNSVRLQSLVLR 48  
 DB 1 RYPSVIMEAKKSHSGCINAE-GKVDHNMNSVTIQEILVL 39

RESULT 8  
 O9TUP4 PRELIMINARY; PRT; 39 AA.

AC O9TUP4:  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE INTERLEUKIN 17 (FRAGMENT).  
 GN IL-17.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Laurent F., Lacroix S.M., Mancassola R., Naciri M.;  
 RT "Partial sequence of ovine IL-17."  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF166488; AAD46377.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 39 AA; 4394 MW; 4A390A22973482A3 CRC64;

Query Match 15.0%; Score 69.5; DB 6; Length 39;  
 Best Local Similarity 42.5%; Pred. NO. 0.13;  
 Matches 17; Conservative 7; Mismatches 15; Indels 1; Gaps 1;

OY 9 RYPOKLAFAECRCGICDARTGRTAALNSVRLQSLVLR 48  
 DB 1 RYPSVIMEAKKSHSGCINAE-GKVDHNMNSVTIQEILVL 39

RESULT 9  
 O9TUP3 PRELIMINARY; PRT; 39 AA.

AC O9TUP3:  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE INTERLEUKIN 17 (FRAGMENT).  
 GN IL-17.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.  
 OX NCBI\_TaxID=9623;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Laurent F., Lacroix S.M., Mancassola R., Naciri M.;  
 RT "Partial sequence of porcine IL-17."

RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF166489; AAD46378.1; -  
FT NON\_TER 1 1  
SQ SEQUENCE 39 AA; 4394 MW; 4A390A22973482A3 CRC64;

Query Match 15.0%; Score 69.5; DB 6; Length 39;  
Best Local Similarity 42.5%; Pred. No. 0.13;  
Matches 17; Conservative 7; Mismatches 15; Indels 1; Gaps 1;

OY 9 RYPOKIAFACLCRCIDARTGRTAALNSVRLQSLVL 48  
111 : 111 : 111 : 111 : 111 : 111 :  
DB 1 RYPSVWEAKSHSGCINAE-GRVDHMHNSVTIQOETVLV 39

RESULT 10  
O9S2C5 PRELIMINARY; PRT; 146 AA.  
AC O9S2C5;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE HYPOTHETICAL 15.0 KDA PROTEIN.  
GN SCF85.16.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Seeger K.J., Harris D.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Redenbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,  
RA Kinashi H., Hopwood D.A.;  
RT "A set of ordered cosmid and a detailed genetic and physical map for  
the 8 Mb Streptomyces coelicolor A3(2) chromosome."  
RL MOL. Microbiol. 21:77-96(1996).  
DR EMBL: AL110470; CAB54173.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 146 AA; 14974 MW; 4350BBBEF04315B6 CRC64;

Query Match 14.5%; Score 67; DB 2; Length 146;  
Best Local Similarity 43.1%; Pred. No. 0.99;  
Matches 22; Conservative 2; Mismatches 15; Indels 12; Gaps 3;

OY 25 IDATGRTAALNSVRLQSLVLRPPCRDGS-GLP--TTCGAFPHTE 71  
111 : 111 : 111 : 111 : 111 : 111 :  
DB 102 LGARTPQTALVGV-ARRPLSHDSAGAPRGRRPFLTRTE 144

RESULT 11  
O13035 PRELIMINARY; PRT; 518 AA.  
AC O13035;  
DT 01-JUL-1997 (TREMBLrel. 04, Created)  
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
DE PROACTIVATOR POLYPEPTIDE PRECURSOR [CONTAINS: SAPOSIN A (PROTEIN A);  
SAPOSIN B (SPHINGOLIPID ACTIVATOR PROTEIN 1) (SAP-1) (DISPERSTIN)  
(SULFATIDE/GMI ACTIVATOR); SAPOSIN C (CO-BETA-GLUCOSIDASE) (AI  
DE ACTIVATOR) (GLUCOSYLKERAMIDASE ACTIVATOR) (SPHINGOLIPID ACTIVATOR  
DE PROTEIN 2) (SAP-2); SAPOSIN D (PROTEIN C) (COMPONENT C)].

GN PSAP.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 194-203.  
RC TISSUE-BRAIN, AND LIVER;  
RX MEDLINE=98129745; Pubmed=9461526;  
RA Azuma N., Seo H.-C., Lie O., Fu Q., Gould R.M., Hiraiwa M., Burt D.W.,  
RA Paton I.R., Morrice D.R., O'Brien J.S., Kishimoto Y.;  
RT "Cloning, expression and map assignment of chicken prosaposin."  
RL Biochem. J. 330:321-327(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Altman N., Horowitz M.;  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: THE LYSOSOMAL DEGRADATION OF SPHINGOLIPIDS TAKES PLACE  
CC BY THE SEQUENTIAL ACTION OF SPECIFIC HYDROLASES. SOME OF THESE  
CC ENZYMES REQUIRE SPECIFIC LOW-MOLECULAR MASS, NON-ENZYMIC PROTEINS:  
CC THE SPHINGOLIPIDS ACTIVATOR PROTEINS (COPROTEINS) (BY SIMILARITY).  
CC -1- FUNCTION: SAPOSIN A AND SAPOSIN C STIMULATE THE HYDROLYSIS OF  
CC GLUCOSYLKERAMIDE BY BETA-GLUCOSYLKERAMIDASE (EC 3.2.1.45) AND  
CC GALACTOSYLKERAMIDE BY BETA-GALACTOSYLKERAMIDASE (EC 3.2.1.46).  
CC SAPOSIN-C APPARENTLY ACTS BY COMBINING WITH THE ENZYME AND ACIDIC  
CC LIPID TO FORM AN ACTIVATED COMPLEX, RATHER THAN BY SUBSTITUTING  
CC THE SUBSTRATE (BY SIMILARITY).  
CC -1- FUNCTION: SAPOSIN B STIMULATES THE HYDROLYSIS OF GALACTO-  
CC CEREBROSIDE SULFATE BY ARYLSULFATASE A (EC 3.1.6.8), GMI  
CC GANGLIOSIDES BY BETA-GALACTOSIDASE (EC 3.2.1.23) AND  
CC GLOBOTRIANOSYLKERAMIDE BY ALPHA-GALACTOSIDASE A (EC 3.2.1.22).  
CC SAPOSIN-B FORMS A SOLUBILIZING COMPLEX WITH THE SUBSTRATES OF THE  
CC SPHINGOLIPID HYDROLASES (BY SIMILARITY).  
CC -1- FUNCTION: SAPOSIN D IS A SPECIFIC SPHINGOMYELIN PHOSPHODIESTERASE  
CC ACTIVATOR (EC 3.1.4.12) (BY SIMILARITY).  
CC -1- SUBUNIT: SAPOSIN B IS A HOMODIMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: LYSOSOMAL (BY SIMILARITY).  
CC -1- PTM: THIS PRECURSOR IS PROTEOLYTICALLY PROCESSED TO 4 SMALL  
CC PEPTIDES, WHICH ARE SIMILAR TO EACH OTHER AND ARE SPHINGOLIPID  
CC HYDROLASE ACTIVATOR PROTEINS (BY SIMILARITY).  
CC -1- PTM: THERE ARE THREE DISULFIDE BRIDGES IN EACH OF THE SAPOSINS.  
CC (BY SIMILARITY).  
CC -1- SIMILARITY: CONTAINS FOUR SAPOSIN/SP-B-B-TYPE DOMAINS AND TWO  
CC A-TYPE DOMAINS. THE B-TYPE DOMAINS CORRESPOND TO THE MATURE  
CC SAPOSINS.  
DR EMBL: AB003471; BA19914.1; -  
DR EMBL: AF108656; AAF05899.1; -  
DR INTERPRO: IPR000004; -  
DR INTERPRO: IPR003119; -  
DR INTERPRO: IPR003259; -  
DR PRAM: PR02199; SAPA; 2.  
DR PRODOM: PD001732; -; 3.  
DR PRODOM: PD012321; -; 1.  
KW Signal; Glycoprotein; Lysosome; Sphingolipid metabolism; Repeat;  
KW GM2-gangliosidosis.  
FT STGNAL 1 17 POTENTIAL.  
FT CHAIN 61 143 SAPOSIN A.  
FT CHAIN 194 276 SAPOSIN B.  
FT CHAIN 307 387 SAPOSIN C.  
FT CHAIN 399 480 SAPOSIN D.  
FT DOMAIN 22 55 SAPOSIN/SP-B A-TYPE DOMAIN.  
FT DOMAIN 485 518 SAPOSIN/SP-B A-TYPE DOMAIN.  
FT CARBOHYD 81 81 POTENTIAL.  
FT CARBOHYD 214 214 POTENTIAL.  
FT CARBOHYD 328 328 POTENTIAL.  
FT CARBOHYD 420 420 POTENTIAL.  
FT CONFLICT 94 94 R -> T (IN REF. 2).  
FT CONFLICT 486 486 E -> D (IN REF. 2).  
SQ SEQUENCE 518 AA; 57601 MW; B803000E891C3563 CRC64;

Query Match 14.0%; Score 65; DB 13; Length 518;



DT	01-Oct-2000 (TReMBLrel. 15, last annotation update)
DE	MEGF8 (FRAGMENT).
GN	MEGF8.
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus
OX	NCBI_TaxID=10116;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	SPRAIN-SPRAGUE-DAWLEY; TISSUE=BRAIN;
RA	Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
RT	"Identification of high-molecular-weight proteins with multiple EGF-
RT	like motifs by motif-trap screening.";
RL	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AB011534; BAA86889.1; -;
DR	HSSP; P35555; IEMN.
DR	INTERPRO; IPR000152; -;
DR	INTERPRO; IPR000561; -;
DR	INTERPRO; IPR001881; -;
DR	INTERPRO; IPR002049; -;
DR	PROSITE; PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR	PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR	PROSITE; PS01865; EGF_2; 2.
DR	PROSITE; PS01248; LAMININ_Type_EGF; UNKNOWN_2.
FT	NON_TER
EQ	SEQUENCE 874 AA: 93764 MW: CB63BDF183DB70AB CRC64;

Query Match	13.28;	Score 61;	DB 11;	Length 874;
Best Local Similarity	27.8%;	Pred. No. 33;		
Matches 20;	Conservative 7;	Mismatches 33;	Indels 12;	Gaps 2;

[illegible]

```
Search completed: February 16, 2001, 12:52:39
Job time: 194 sec
```



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2001, 12:49:50 ; Search time 14.73 Seconds

(without alignments)  
190.739 Million cell updates/sec

Title: US-09-320-713-2

Perfect score: 463

Sequence: 1 ARVDDEDEHPQKLAFAECL.....FHTEFHPVGCCTVLPKRV 87

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues

Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	125	27.0	155	1 IL17_HUMAN	Q16552 homo sapien
2	124	26.8	151	1 VG13_HSVSA	P24916 herpesvirus
3	119	25.7	150	1 IL17_RAT	O61453 rattus norv
4	119	25.7	158	1 IL17_MOUSE	O62386 mus musculu
5	63	13.6	362	1 A2HS_PIG	P28700 sus scrofa
6	62.5	13.5	762	1 P115_CHICK	Q98917 gallus gall
7	60.5	13.1	306	1 YE16_YEAST	P40101 saccharomyc
8	60.5	13.1	909	1 CT1A_FUSO	P52958 fusarium so
9	60	13.0	344	1 VD33_MYCTU	O10644 mycobacteri
10	60	13.0	454	1 DAT_HAEIN	P44951 haemophilus
11	58.5	12.6	640	1 CLAT_PIG	P13222 sus scrofa
12	57.5	12.4	452	1 CLPX_HELPJ	O92118 heliobacte
13	57	12.3	303	1 YJ85_MYCTU	O10872 mycobacteri
14	56.5	12.2	455	1 ARQA_AGREP	O94462 agrobacteri
15	56	12.1	398	1 CAR4_RHINI	O03700 rhizopus ni
16	56	12.1	798	1 YLEFN_CAEEL	Q20296 rhizopus ni
17	36	12.1	1964	1 NTC4_MOUSE	P31695 mus musculu
18	55.5	12.0	298	1 M1AA_AGRU	P34436 agrobacteri
19	55.5	12.0	328	1 PIT_BACSV	O34366 bacillus su
20	55.5	12.0	363	1 DIDR_CARAU	P35406 carassius su
21	55.5	12.0	445	1 DAT_CARAU	P56744 actinobact
22	55.5	12.0	547	1 PRB_BOMMO	P28523 bombyx mori
23	35	11.9	350	1 REDD_STKCO	P16922 streptomyce
24	55	11.9	524	1 SAP_HUMAN	O07602 h piractinva
25	55	11.9	635	1 GIDA_SYNY3	O56994 synecocyst
26	54.5	11.8	470	1 RBSA_RHME	O923r2 rhizobium m
27	54.5	11.8	491	1 ACHE_BOVIN	P02715 bos taurus
28	54.5	11.8	493	1 ACHE_HUMAN	O04844 homo sapien
29	54.5	11.8	530	1 IDS_HUMAN	P22304 homo sapien
30	54.5	11.8	782	1 ZFY1_MOUSE	P10925 mus musculu
31	54.5	11.8	783	1 ZFY2_MOUSE	P20662 mus musculu
32	54	11.7	305	1 Y441_TREPA	O84455 treponema p
33	53.5	11.6	446	1 CLPX_HELPJ	O25926 heliobacte

34	53.5	11.6	499	1 KPY1_TRYBB	P30615 trypanosoma
35	53.5	11.6	499	1 KPY2_TRYBB	P30616 trypanosoma
36	53.5	11.6	1587	1 TOP2_PENCH	O29898 penicillium
37	53.5	11.6	1687	1 Z142_HUMAN	P2746 homo sapien
38	53	11.4	412	1 NH64_CAEEL	O44960 caenorhabdi
39	53	11.4	456	1 YRDX_RHOSH	O01856 rhodobacter
40	53	11.4	3051	1 YNXX_CAEEL	P34576 caenorhabdi
41	53	11.4	5032	1 RYNR_HUMAN	P21817 homo sapien
42	52.5	11.3	278	1 MA32_MOUSE	O35658 mus musculu
43	52.5	11.3	495	1 CD5_HUMAN	P06127 homo sapien
44	52.5	11.3	495	1 HM18_CAEEL	P41934 caenorhabdi
45	52.5	11.3	785	1 VD05_VACCC	P21010 vaccinia vi

## ALIGNMENTS

RESULT	ID	IL17_HUMAN	STANDARD:	PRT:	155 AA.
AC	Q16552:				
DT	01-NOV-1997	(Rel. 35, Created)			
DT	01-NOV-1997	(Rel. 35, Last sequence update)			
DE	15-DEC-1998	(Rel. 37, Last annotation update)			
DE	INTERLEUKIN-17 PRECURSOR (IL-17) (CYTOTOXIC T LYMPHOCYTE-ASSOCIATED				
DE	ANTIGEN 8) (CTLA-8).				
GN	IL17 OR CTLA8.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
CC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=96281911; PubMed=8676080;				
RA	Fossiez F., Djossou O., Chomarat P., Flores-Romo L., Alt-Yahia S.,				
RA	Maat C., Pin J.-J., Garrone P., Garcia E., Saeland S., Blanchard D.,				
RA	Gallard C., Das Mahapatra B., Rouvier E., Golstein P., Banchereau J.,				
RA	Lebecque S.;				
RT	"T cell Interleukin-17 induces stromal cells to produce				
RT	proinflammatory and hematopoietic cytokines."				
RL	J. Exp. Med. 183:2593-2603(1996).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=96094436; PubMed=7499828;				
RA	Yao Z., Painter S.L., Fanslow W.C., Ulrich D., Macduff B.M.,				
RA	Spriggs M.K., Armitage R.J.;				
RT	"Human IL-17: a novel cytokine derived from T cells."				
RL	J. Immunol. 155:5483-5486(1995).				
CC	- FUNCTION: INDUCES STROMAL CELLS TO PRODUCE PROINFLAMMATORY AND				
CC	HEMATOPOIETIC CYTOKINES. ENHANCES THE SURFACE EXPRESSION OF THE				
CC	INTRACELLULAR ADHESION MOLECULE-1 (ICAM-1) IN FIBROBLASTS.				
CC	- SUBUNIT: HOMODIMER.				
CC	- TISSUE SPECIFICITY: RESTRICTED TO ACTIVATED MEMORY T-CELLS.				
CC	- FUNCTION: FOUND BOTH IN GLYCOSTATED AND NONGLYCOSTATED FORMS.				
CC	- SIMILARITY: STRONG, TO OTHER MAMMALIAN INTERLEUKIN-17 AND TO				
CC	HERPESVIRUS SAIMIRI IMMEDIATE EARLY GENE 13 PROTEIN.				
CC	-----				
CC	THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION				
CC	BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTSTATION -				
CC	THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS				
CC	USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY				
CC	MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL				
CC	ENTITIES REQUIRES A LICENSE AGREEMENT (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	OR SEND AN EMAIL TO <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).				
DR	EMBL; 258820; CAAG1233.1; -				
DR	EMBL; U32659; AAC50341.1; -				
DR	MM; 603149; -				
FT	CYCLOKINE; Glycoprotein; T-cell; Antigen; Signal.				
FT	SIGNAL	1	23	POTENTIAL.	
FT	CHAIN	24	155	INTERLEUKIN-17.	
FT	CARBOHYD	68	68	N-LINKED (GLCNAC. . .) (POTENTIAL).	







RESULT 8  
CTIA\_FUSO STANDARD: PRT: 909 AA.  
ID CTIA\_FUSO  
AC P52958;  
DR 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DE 15-JUL-1999 (Rel. 38, Last annotation update)  
GN CUTINASE TRANSCRIPTION FACTOR 1 ALPHA.  
OS Fusarium solani (subsp. pisi) (Nectria haematococca).  
OC Eukaryota; Fungi; Ascomycota; anamorphic Ascomycota; Fusarium.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Li D., Kolattukudy P.E.;  
RL Submitted (MAR-1996) to the EMBL/Genbank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (POTENTIAL).  
CC -1- SIMILARITY: CONTAINS A ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR CLUSTER DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: U51671; AAA96824.1; -  
DR INTERPRO: IPR001138; -  
DR PFAM: PF00172; ZN\_clus: 1.  
DR PROSITE: PS00463; ZN2\_CYS6\_FUNGAL\_1; 1.  
DR PROSITE: PS00463; ZN2\_CYS6\_FUNGAL\_2; 1.  
KW Transcription regulation; DNA-binding; Nuclear protein; Zinc; Metal-binding.  
FT ZN-BIND 61 90 ZN(2)-CYS(6), FUNGAL-TYPE.  
FT DNA-BIND 61 90 ZN(2)-CYS(6), FUNGAL-TYPE.  
SQ SEQUENCE 909 AA; 101172 MW; 66AC2EC18F5C96A9 CRC64;  
  
Query Match 13.1%; Score 60.5; DB 1; Length 909;  
Best Local Similarity 28.9%; Pred. No. 13;  
Matches 24; Conservative 8; Mismatches 26; Indels 25; Gaps 3;  
  
OY 12 OKLAF---AECRCRCIDA-----FTGRTALNSVRLQSLVLR 50  
DB 240 QGAFLLPRLACDELIDYFMSVHPYVINTREMPROYRDKPKPSILLQSVLLACT 299  
OY 51 RPSCSR---DGSGLPYPGAFPH 69  
DB 300 RACNNALMDANGSTTPALTRY 322  
  
RESULT 9  
YD33\_MYCTU STANDARD: PRT: 344 AA.  
ID YD33\_MYCTU  
AC Q10644;  
DR 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DE 30-MAY-2000 (Rel. 39, Last annotation update)  
GN HYPOTHETICAL 33.9 KDA PROTEIN RV1333.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Bacteria; Firmicutes; Actinobacteriae; Mycobacteriaceae; Mycobacterium.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN-H37RV;  
RA MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D., Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F., Badcock K., Basham D., Brown D., Chillingworth T., Connor R., Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,

RA Hornsby T., Jagels K., Krogh A., Mclean J., Moule S., Murphy L., Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J., Rutter S., Seeger K., Skelton S., Squares S., Squires R., Sulston J., Taylor K., Whitehead S., Barrett B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence."  
RL Nature 393:537-544(1998).  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -1- SIMILARITY: STRONG, TO M.LEPRAE B1549\_C2\_208.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: Z73902; CAA98097.1; -  
DR MEROPS: T04.001; -  
DR TUBERCULIST; RV1333; -  
DR Hypothetical protein; Transmembrane.  
FT TRANSMEM 25 45 POTENTIAL.  
FT TRANSMEM 68 88 POTENTIAL.  
FT TRANSMEM 104 124 POTENTIAL.  
FT TRANSMEM 133 153 POTENTIAL.  
FT TRANSMEM 161 181 POTENTIAL.  
FT TRANSMEM 224 244 POTENTIAL.  
FT TRANSMEM 276 296 POTENTIAL.  
FT TRANSMEM 302 322 POTENTIAL.  
SQ SEQUENCE 344 AA; 33952 MW; 296A055FECF8B650 CRC64;  
  
Query Match 13.0%; Score 60; DB 1; Length 344;  
Best Local Similarity 32.5%; Pred. No. 5.1;  
Matches 27; Conservative 7; Mismatches 17; Indels 32; Gaps 6;  
  
OY 23 GCIDARTG---RENAAL---NSVRLQSL-----VLR-----RRCSD 56  
DB 43 GAVDGRGAPGTRETDLPANSVRFVALLAGSAYGLAADGVMLERHGRVAMD 102  
OY 57 GSGLP-TPGAFPHTEFHVPG 78  
DB 103 SGVPIVPGAVIF-----DLPVG 120  
  
RESULT 10  
DAT\_HAEIN STANDARD: PRT: 454 AA.  
ID DAT\_HAEIN  
AC P44951;  
DR 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DE 30-MAY-2000 (Rel. 39, Last annotation update)  
GN DIAMINOBUTYRATE-PYRUVATE AMINOTRANSFERASE (EC 2.6.1.46) (L-DIAMINOBUTYRIC ACID TRANSAMINASE) (DIAMINOBUTYRATE TRANSAMINASE) (DABA AMINOTRANSFERASE) (DABA-AT).  
DE DAT OR H10949.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae; Haemophilus.  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN-RD / KW20;  
RA MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M., McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D., Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M., Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D., Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C., Fier L.D., Fitchman J.L., Furman J.L., Geoghegan N.S.M., Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O., Venter J.C.;

```

RT "Whole-genome random sequencing and assembly of Haemophilus
RT Influenzae Rd."
RL Science 269:496-512(1995).
CC -1- CATALYTIC ACTIVITY: L-2,4-DIAMINOBUTANOATE + 2-OXOGLOUTARATE = L-
CC ASPARTATE 4-SEMIALDEHYDE + L-GLUTAMATE.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (POTENTIAL)
CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U33776; AAC22610.1; -.
DR HSSP: P16932; IDGE.
DR TIGR: H10949; -.
DR INTERPRO: IPR000954; -.
DR PFAM: PF00202; aminotran.3; 1.
DR PROSITE: PS00600; AA-TRANSFER CLASS.3; 1.
KW Transferase: Aminotransferase, Pyridoxal phosphate.
FT BINDING 287 287 PYRIDOXAL PHOSPHATE (POTENTIAL).
SQ SEQUENCE 454 AA; 49368 MW; B4B26F3AB7C55063 CRC64;

Query Match 13.0%; Score 60; DB 1; Length 454;
Best Local Similarity 30.6%; Pred. No. 6.8;
Matches 26; Conservative 8; Mismatches 21; Indels 30; Gaps 5;

QY 10 YPKLIAFAECLRCG-----IDARTGREPAL--NSVRLQSLVLRRCPSRD- 56
DQ 31 YPKLPIFAVAKAGCGWTVDEGNEYLDPLAGAGTLALGHNPIILMQAL-----KDV 81
QY 57 -GSGLP-----TPGAFATFEFT 73
DQ 82 LDGSLPLHTLDTPTPLKDAFSEELL 106

RESULT 11
CLAT_PIG STANDARD; PRT; 640 AA.
AC P13222;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE CHOLINE O-ACETYLTRANSFERASE (EC 2.3.1.6) (CHOLACTASE) (CHOLINE
DE ACETYLASE).
GN CHAT.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Cetartiodactyla; Suidae; Sus.
CC [1]
RN SEQUENCE FROM N.A.
RC TISSUE=VENTRAL SPINAL CORD;
RX MEDLINE=80097472; PubMed=3480542;
RA Berrard S., Brice A., Lottspeich F., Braun A., Barde Y.-A., Mallet J.;
RT "CDNA cloning and complete sequence of porcine choline
RT acetyltransferase: in vitro translation of the corresponding RNA
RT yields an active protein."
RL Proc. Natl. Acad. Sci. U.S.A. 84:9280-9284(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=VENTRAL SPINAL CORD;
RX MEDLINE=89229974; PubMed=2713713;
RA Berrard S., Brice A., Mallet J.;
RT "Molecular genetic approach to the study of mammalian choline
RT acetyltransferase."
RL Brain Res. Bull. 22:147-153(1989).
RN [3]
RP SEQUENCE OF 1-11.

```

```

RC TISSUE=BRAIN;
RX MEDLINE=87085562; PubMed=3794697;
RA Braun A., Barde Y.-A., Lottspeich F., Mewes H.-W., Thoenen H.;
RT "N-terminal sequence of pig brain choline acetyltransferase purified
RT by a rapid procedure."
RL J. Neurochem. 48:16-21(1987)
CC -1- FUNCTION: CATALYZES THE BIOSYNTHESIS OF THE NEUROTRANSMITTER
CC ACETYLCHOLINE.
CC -1- CATALYTIC ACTIVITY: ACETYL-COA + CHOLINE = COA + O-ACETYLCHOLINE.
CC -1- SIMILARITY: BELONGS TO THE CARNITINE/CHOLINE ACETYLTRANSFERASE
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J03021; AAA31000.1; -.
DR EMBL: M27736; AAA31015.1; -.
DR PIR: A28047; A28047.
DR PIR: A39961; A39961.
DR INTERPRO: IPR000542; -.
DR PFAM: PF00755; Carn_acetyltransf.1.
DR PROSITE: PS00439; ACYLTRANSF.C.1; 1.
DR PROSITE: PS00440; ACYLTRANSF.C.2; 1.
KW Transferase: Acyltransferase; Neurotransmitter biosynthesis.
FT INT_MET 0 0 POTENTIAL.
FT ACT_SITE 333 333
SQ SEQUENCE 640 AA; 71599 MW; 5EC27BE8B7CC317 CRC64;

Query Match 12.6%; Score 58.5; DB 1; Length 640;
Best Local Similarity 32.5%; Pred. No. 15;
Matches 13; Conservative 9; Mismatches 13; Indels 5; Gaps 1;

QY 19 CLCRCIDARTGREPALNSVRLQSLVLRRCPSRDG 58
DQ 276 CILVCLDAPGCMELSDTN-----RALQLHGCGCSKNGA 310

RESULT 12
CLPX_HELPJ STANDARD; PRT; 452 AA.
AC Q9JLJ8;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE ATP-DEPENDENT CLP PROTEASE ATP-BINDING SUBUNIT CLPX.
GN CLPX OR HP1374.
OS Helicobacter pylori j99 (Campylobacter pylori j99).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
CC [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=99120557; PubMed=9923682;
RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Doig P.C.,
RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
RA Tummino P.J., Caruso A., Uria-Nickelsen M., Mills D.W., Ives C.,
RA Gibson R., Merberg D., Mills S.D., Jiang Q., Taylor D.E., Vovis G.F.,
RA Trust T.J.;
RT "Genomic sequence comparison of two unrelated isolates of the human
RT gastric pathogen Helicobacter pylori."
RL Nature 387:176-180(1999).
CC -1- FUNCTION: ATP-DEPENDENT SPECIFICITY COMPONENT OF THE CLP PROTEASE.
CC IT DIRECTS THE PROTEASE TO SPECIFIC SUBSTRATES. CAN PERFORM
CC CHAPERONE FUNCTIONS IN THE ABSENCE OF CLPX (BY SIMILARITY).
CC -1- SUBUNIT: HETERODIMER OF CLPX AND CLPX (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CLPX CHAPERONE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AE001551; AAD06862.1; -  
DR INTERPRO: IPR001939; -  
DR PIRAM: PF00004; AAA; 1.  
DR Chaperone; ATP-binding.  
FT ZN\_FING 7 39 C4-TYPE.  
FT NP\_BIND 155 162 ATP (POTENTIAL).  
SQ SEQUENCE 452 AA; 51166 MW; 3C1DA719296474B CRC64;

Query Match 12.4%; Score 57.5; DB 1; Length 452;  
Best Local Similarity 29.3%; Pred. No. 14;  
Matches 22; Conservative 8; Mismatches 30; Indels 15; Gaps 3;

OY 7 EDRYOK--LAF--CICRCIDARTGRTALNSVLLSLVLRPPCSRDGS 58  
DB 14 ESRDRKRRIIPASINMKDVCCEYCIDVHGLHLYDSKYDRMDSLALKDRRLRMES 73

OY 59 GLPTGAFATFTEFI 73  
DB 74 S-----AYEEFL 81

RESULT 13  
Y85 MYCTU STANDARD: PRT: 303 AA.  
AC 010872; -

DT 01-OCT-1996 (Rel. 34, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE HYPOTHETICAL TRANSCRIPTIONAL REGULATOR RV1985C.  
GN RV1985C OR MTCY39.34.

OS Bacterium tuberculosis.  
OC Bacteria; Filumetes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.

RA [1]  
RC SEQUENCE FROM N.A.

RA MEDLINE-96295987; PubMed-9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jags K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Ruter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E.,  
RA Taylor K., Whitehead S., Barrett B.G.;  
RA "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence.";  
RL Nature 393:537-544(1998).

CC -1- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL  
CC REGULATIONS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: Z74025; CA98410.1; -  
DR TUBERCULIST; RV1985C; -  
DR INTERPRO: IPR000847; -  
DR PIRAM: PF00126; HTH\_L; 1.  
DR PRINTS: PR00039; HTHLYSR.  
DR PROSITE: PS00044; HTH\_LYSR\_FAMILY; 1.

KW Hypothetical protein; Transcription regulation; DNA-binding.  
FT DNA\_BIND 23 42 H-T-H MOTIF (POTENTIAL).  
SQ SEQUENCE 303 AA; 32836 MW; 344BDB1056207144 CRC64;

Query Match 12.3%; Score 57; DB 1; Length 303;  
Best Local Similarity 28.8%; Pred. No. 10;  
Matches 17; Conservative 10; Mismatches 24; Indels 8; Gaps 2;

OY 11 POKLAFECICRCIDARTGR---ETALNSVRLQ---SLVLRPPCSRDGSLP 61  
DB 9 POLAALAAVEELGSDPAERLHVTPSAVSQRKSLQGVGVVVRKPCRTATGAP 67

RESULT 14  
ARO\_AGRSP STANDARD: PRT: 455 AA.  
AC 094E4; -

DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 01-OCT-2000 (Rel. 40, Last annotation update)  
DE 3-PHOSPHOSHIKIMATE 1-CARBOXYVINYLTRANSFERASE (EC 2.5.1.19) (5-  
DE ENOLPYRUVYLSHIKIMATE-3-PHOSPHATE SYNTHASE) (EPSP SYNTHASE) (EPSPS).  
GN AROA.

OS Agrobacterium sp. (strain CP4).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Agrobacterium.

RA [1]  
RC SEQUENCE FROM N.A., AND SEQUENCE OF 1-28; 47-61 AND 321-333.  
RA Barry G.F., Kishore G.M., Padgett S.R., Stallings W.C.;  
RT "Glucose-tolerant 5-enolpyruvylshikimate-3-phosphate synthases";  
RL Patent number US5633435, 27-MAY-1997.

RA [2]  
RC SEQUENCE OF 1-15.  
RA MEDLINE-96182485; PubMed-8598558;

RA Harrison L.A., Bailey M.R., Naylor M.W., Ream J.E., Hammond B.G.,  
RA Nida D.L., Bunette B.L., Nickson T.E., Miteky T.A., Taylor M.L.,  
RA Fuchs R.L., Padgett S.R.;

RA "The expressed protein in glyoxylate-tolerant soybean, 5-  
RT enolpyruvylshikimate-3-phosphate synthase from Agrobacterium sp.  
RT strain CP4, is rapidly digested in vitro and is not toxic to acutely  
RT gavaged mice.";  
RL J. Nutr. 126:728-740(1996).

CC -1- CATALYTIC ACTIVITY: PHOSPHOENOLPYRUVATE + 3-PHOSPHOSHIKIMATE -  
CC ORTHOPHOSPHATE + O(5)-(1-CARBOXYVINYL)-3-PHOSPHOSHIKIMATE.

CC -1- PATHWAY: SIXTH STEP IN THE BIOSYNTHESIS OF CHORISMATE WITHIN  
CC THE BIOSYNTHESIS OF AROMATIC AMINO ACIDS (THE SHIKIMATE PATHWAY).

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).  
CC -1- BIOTECHNOLOGY: INTRODUCED BY GENETIC MANIPULATION AND EXPRESSED IN  
CC GLYCOLATE-TOLERANT SOYBEAN, CANTOLA, COTTON AND MAIZE BY MONSANTO.

CC DEVELOPED TO PROVIDE NEW WEED-CONTROL OPTIONS FOR FARMERS.

CC EXPRESSION OF THIS PROTEIN IN PLANTS IMPARTS HIGH LEVELS OF  
CC GLYCOLATE TOLERANCE.

CC -1- SIMILARITY: BELONGS TO THE EPSP SYNTHASE FAMILY.  
DR PROSITE: PS00104; EPSP SYNTHASE\_1; 1.  
DR PROSITE: PS00885; EPSP SYNTHASE\_2; 1.

KW Aromatic amino acid biosynthesis; Transferase; Herbicide resistance;  
KW Genetically modified food.  
FT CONFLICT 2 2 S -> L (IN REF. 2).  
SQ SEQUENCE 455 AA; 47588 MW; 236580D08D6EF422 CRC64;

Query Match 12.2%; Score 56.5; DB 1; Length 455;  
Best Local Similarity 28.7%; Pred. No. 18;  
Matches 27; Conservative 10; Mismatches 32; Indels 25; Gaps 5;

OY 3 VOTDEDRYPOKL-----AF--CICRCIDAR---GRTALNSVRLQ----- 42  
DB 314 VTPVEDRAPASMTDEYPIILAAVAFAGATVWNGLELRFVESDRUSAAVNGLKGVCDD 373

OY 43 --QSLVLRPPCSRDGSLPTGAFATFTEFIH 74  
DB 374 EGETSLVVRGRP---DGKGLGNASAAVAATHD 404

```

RESULT 15
CAR4_RHINI STANDARD; PRT; 398 AA.
AC 003700;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE RHIZOPUSPEPSIN 4 PRECURSOR (EC 3.4.23.21) (ASPARTATE PROTEASE).
OS Rhizopus niveus.
OC Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC Rhizopus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YAMAZAKI / IFO 4810;
RA Horiiuchi H., Nakamura H., Okazaki T., Yano K., Takagi M.;
RL Submitted (DEC-1990) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF PROTEINS WITH BROAD SPECIFICITY
CC SIMILAR TO THAT OF PEPsin A, PREFERING HYDROPHOBIC RESIDUES AT P1
CC AND P1'. CLOTS MILK AND ACTIVATES TRYPSINOGEN. DOES NOT CLEAVE
CC 4-GLN-1-HIS-5, BUT DOES CLEAVE 10-HIS-1-LEU-11 AND 12-VAL-1-GLU-13
CC IN B CHAIN OF INSULIN.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1; ALSO KNOWN AS THE
CC EUKARYOTIC ASPARTYL PROTEASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X56992; CAA40309.1; .
DR HSP: P06026; 4APR.
DR MEROPS: A01.012; -.
DR INTERPRO: IPR001461; -.
DR INTERPRO: IPR001969; -.
DR PFAM: PF000026; asp. 1.
DR PRINTS: PR00792; PEPsin.
DR PROSITE: PS00141; ASP_PROT3ASE; 2.
KW Hydrolase; Aspartyl protease; zymogen; Signal; Multigene family.
FT SIGNAL 1 ? ?
FT PROPEP 74 74 POTENTIAL.
FT CHAIN 75 398 RHIZOPUSPEPSIN 4.
FT ACT_SITE 108 108 BY SIMILARITY.
FT ACT_SITE 291 291 BY SIMILARITY.
FT DISULFID 121 124 BY SIMILARITY.
FT DISULFID 325 358 BY SIMILARITY.
FT DOMAIN 58 67 POLY-SER.
FT DOMAIN 237 240 POLY-GLY.
SQ SEQUENCE 398 AA; 41408 MW; 70CA63FEFBC05D5 CRC64;

Query Match 12.1%; Score 56; DB 1; Length 398;
Best Local Similarity 22.3%; Pred. No. 18;
Matches 21; Conservative 11; Mismatches 30; Indels 32; Gaps 3;

QY 2 RVDTDERTYPOKLAFAEGLCRGC-----IDART-----G 30
   : : | | | | | | | | | | | | | | | | | | | |
Db 102 KLIKDFPTGSSDLWFASTLCITNCGSSQTKYDPSQSYAKDGRWTSISYGDSSASGILG 161
   : : | | | | | | | | | | | | | | | | | | | |
OY 31 RETAALNSVRLQSLVLRPPGSRDGSGLPTPG 64
   : : | | | | | | | | | | | | | | | | | | | |
Db 162 KDTVNLGLGLIKNOITIELAKREASSFSSG-PSDG 194
   : : | | | | | | | | | | | | | | | | | | | |

```

Search completed: February 16, 2001, 12:52:57  
 Job time: 187 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2001, 12:48:14 ; Search time 23.57 Seconds  
(without alignments)  
250.630 Million cell updates/sec

Title:

US-09-320-713-2

Perfect score:

463

Sequence: 1 ARVDTEDEHYPPQKLAFAECL.....FHTEFIHPVGCCTVLPKRV 87

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	124	26.8	151	1 B45351	immediate-early pr
2	119	25.7	147	2 JC4628	cytotoxic T-lympho
3	119	25.7	150	2 I49623	cytotoxic T-lympho
4	88	19.0	148	2 T21334	hypothetical prote
5	67	14.5	146	2 T36476	hypothetical prote
6	63.5	13.7	563	2 T04359	pectin methylster
7	63	13.6	362	2 S22395	feculin precursor
8	62	13.4	1737	2 T00209	MEGF8 protein - hu
9	61	13.2	2222	2 T13924	sdh protein - frui
10	60.5	13.1	306	2 S50689	hypothetical prote
11	60	13.0	227	2 I39313	zinc-finger protei
12	60	13.0	245	2 T27780	hypothetical prote
13	60	13.0	344	2 A70771	probable hydrolase
14	60	13.0	436	2 T36490	gamma-aminobutyrat
15	60	13.0	454	2 C64104	hypothetical prote
16	59.5	12.9	491	2 T19635	hypothetical prote
17	59.5	12.9	581	2 T04844	probable serine/th
18	58.5	12.6	102	2 T36634	probable transcrip
19	58.5	12.6	641	2 A39961	cholesterol O-acetyl
20	58.5	12.6	907	2 T02417	probable finger pr
21	58	12.5	183	2 A75605	hypothetical prote
22	58	12.5	194	2 T00941	hypothetical prote
23	58	12.5	524	2 E75574	aldehyde dehydroge
24	58	12.5	4767	2 T31345	hypothetical prote
25	57.5	12.4	452	2 F71826	ATP-dependent prot
26	57.5	12.4	877	2 T43449	hypothetical prote
27	57.5	12.4	1067	2 T04274	hypothetical prote
28	57	12.3	303	2 G70756	hypothetical prote
29	56.5	12.2	430	2 T12541	hypothetical prote

30	56	12.1	204	2 T28939	hypothetical prote
31	56	12.1	1964	2 T09059	notch4 - mouse
32	55.5	12.0	298	2 B42643	tRNA isopentenyltr
33	55.5	12.0	328	2 A69855	low-affinity inorg
34	55.5	12.0	363	2 I50475	dopamine D1 recept
35	55.5	12.0	375	2 JC7287	G-protein coupled
36	55.5	12.0	445	2 T43804	gamma-aminobutyrat
37	55.5	12.0	489	2 S68280	protein disulfide-
38	55.5	12.0	547	2 S19607	alkaline phosphata
39	55.5	12.0	534	2 T33894	hypothetical prote
40	55.5	12.0	597	2 T21883	hypothetical prote
41	55	11.9	297	2 J10032	hypothetical 31.7k
42	55	11.9	350	2 T34829	transcription regu
43	55	11.9	456	2 S76369	hypothetical prote
44	55	11.9	527	1 SAHUP	saposin precursor
45	55	11.9	634	2 A83386	hypothetical prote

#### ALIGNMENTS

```
RESULT 1
B45351
Immediate-early protein 2 - saimiriine herpesvirus 1 (strain 11)
N:Alternate names: hypothetical protein ORF13
C:Species: saimiriine herpesvirus 1
A:Note: host Saimiri sciureus (common squirrel monkey)
C:Date: 30-Sep-1993 #sequence_textvison 30-Sep-1993 #text_change 16-Jul-1999
C:Accession: B45351; D36807
R:Nicholas, J.; Smith, E.P.; Coles, L.; Honess, R.
Virology 179, 189-200, 1990
A:Title: Gene expression in cells infected with gammaherpesvirus saimiri: properties
A:Reference number: A45351; MUID:91021021
A:Accession: B45351
A:Molecule type: mRNA
A:Residues: 1-151 <NIC>
A:Cross-references: GB:M6286; NID:g331040; PIDN:AAA46156.1; PID:g331042
R:Albrecht, J.
Submitted to the EMBL Data Library, January 1992
A:Description: Primary structure of the herpesvirus saimiri genome.
A:Reference number: A36806
A:Accession: D36807
A:Molecule type: DNA
A:Residues: 1-151 <ALB>
A:Cross-references: GB:X64346; NID:g60320; PIDN:CAA45636.1; PID:g60334
R:Albrecht, J.C.; Nicholas, J.; Biller, D.; Cameron, K.R.; Biesinger, B.; Newman, C.;
J. Virol. 66, 5047-5058, 1992
A:Title: Primary structure of the herpesvirus saimiri genome.
A:Reference number: A37309; MUID:92333688
A:Contents: annotation; protein-coding frames
A:Note: neither protein nor nucleotide sequence is given
C:Genetics:
A:Gene: 13
C:Superfamily: saimiri herpesvirus immediate-early protein 2
C:Keywords: early protein

Query Match 26.8%; Score 124; DB 1; Length 151;
Best Local Similarity 37.0%; Pred. No. 1.6e-07;
Matches 30; Conservative 13; Mismatches 24; Indels 14; Gaps 3;

OY 6 DEDRPPQKLAFAECLRCRCIDARTGRTALNSVRLDLSLVLR--RPCSRDGSGLPTTP 63
:||||| :| :||| :||| :||| :||| :||| :|||
Db 77 DODRPPSVIEWKCKRYLCCVNA-DGNVDYHMNSVPIQDEIIVVRKHOPCPN----- 127
OY 64 GAFAHTEFIHPVGCCTVLP 84
:| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 128 ---STRLEKMLVTVCCTVTP 145

RESULT 2
JC4628
cytotoxic T-lymphocyte-associated antigen 8 precursor - mouse
```





Db 81 P--LFAFRTK 88

RESULT 11

I39313

zinc-finger protein (ZNFPT1) - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 06-Sep-1996 #sequence\_revision 06-Sep-1996 #text\_change 05-Nov-1999

C/Accession: I39313

R:Heuber, K.; Druck, T.; LaForgia, S.; Lasota, J.; Croce, C.M.; Lanfranccone, L.; Dotti,

Hum. Genet. 91, 217-222, 1993

A>Title: Chromosomal localization of four human zinc finger CDNAs.

A:Reference number: I39313; MUID:93239177

A/Accession: I39313

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-227 <RES>

A:Cross-references: EMBL:X65230; NID:g505543; PIDN:CAA46337.1; PID:g505544

Query Match 13.0%; Score 60; DB 2; Length 227;

Best Local Similarity 29.7%; Pred. No. 10;

Matches 19; Conservative 6; Mismatches 31; Indels 8; Gaps 2;

Y 11 POKLAFECICRGICDARTGRTALNSVRLQSLVLRPPCSRDGSLPTPGAFAPHT 70

Db 87 PMKTQVKEPC-NCODSRTA-----SVRYKSLSSKSKYECQCKGKAFICPSSFRGHV 138

Y 71 EFH 74

Db 139 NSHH 142

RESULT 12

T27780

hypothetical protein ZK218.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 26-May-2000

C/Accession: T27780

R:McMurray, A.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z20418

A/Accession: T27780

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-245 <WTL>

A:Cross-references: EMBL:Z82085; PIDN:CA804984.1; GSPDB:GN00023; CESP:ZK218.1

A:Experimental source: clone ZK218

C:Genetics:

A:Gene: CESP:ZK218.1

A:Map position: 5

A:Introns: 35/3; 124/1

C:Superfamily: Caenorhabditis elegans hypothetical protein C4967.3

Query Match 13.0%; Score 60; DB 2; Length 245;

Best Local Similarity 29.7%; Pred. No. 11;

Matches 19; Conservative 8; Mismatches 21; Indels 16; Gaps 3;

Y 19 CLRGCDIARTG--RETALNSVRLQSLVLRPPCSRDGSLPTPGAFAPHTFTHVP 76

Db 157 CNAGCGVDIAIMCANDISICNNVGMQDFVNLNCQRTCSRCSSA-PVPG----- 203

Y 77 VGC 80

Db 204 -GCT 206

RESULT 13

A70771

probable hydrolase - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C/Accession: A70771

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd,

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulton, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:9825987

A/Accession: A70771

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-344 <COL>

A:Cross-references: GB:T73902; GB:AL123456; NID:g3261576; PIDN:CAA98097.1; PID:g4502

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: RV1333

Query Match 13.0%; Score 60; DB 2; Length 344;

Best Local Similarity 32.5%; Pred. No. 16;

Matches 27; Conservative 7; Mismatches 17; Indels 32; Gaps 6;

Y 23 GCDIARTG----RETAL--NSVRLQSL-----VLR-----RPPCSRD 56

Db 43 GAVDCRGAPGRTRETDLDPANSVRFYDALAGSAYGLAADGVKWLDEHRRGVAMD 102

Y 57 GSGLP-TPGAFAPHTFTHVPV 78

Db 103 SGVPIVPGAVIF-----DLPVG 120

RESULT 14

T36490

probable transmembrane transport protein - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000

C/Accession: T36490

R:Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M

submitted to the EMBL Data Library, July 1999

A:Reference number: Z21608

A/Accession: T36490

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-436 <SAU>

A:Cross-references: EMBL:AL096822; PIDN:CA846932.1; GSPDB:GN00070; SCOEDB:SCGD3.10C

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SCGD3.10C

C:Superfamily: bicyclomycin resistance protein

Query Match 13.0%; Score 60; DB 2; Length 436;

Best Local Similarity 36.2%; Pred. No. 20;

Matches 21; Conservative 7; Mismatches 20; Indels 10; Gaps 2;

Y 25 IDARTGRTALNSVRLQSL-----LVLRPPCSRDGSLPTPGAFAPHTFTHVPV 77

Db 384 LDATGDDYTVAFSSVFYQALGVQILRLRKRALRRERL-----VASRVEYTHVPV 436

RESULT 15

C64104

gamma-aminobutyrate aminotransferase (EC 2.6.1.-) - Haemophilus influenzae (strain Rd

C:Species: Haemophilus influenzae

C>Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 08-Oct-1999

C/Accession: C64104; JC5879

R:Flieschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerley, J.

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman

, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Georgagen, N.B.M.

Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Vente

A>Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A:Reference number: A64000; MUID:95350630



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2001, 12:48:00 ; Search time 18.64 Seconds  
(without alignments)  
83.812 Million cell updates/sec

Title: US-09-320-713-2

Perfect score: 463  
Sequence: 1 ARVDTEDEPQKLAFAECL.....FHTFEIHVPVGCIVLPRSV 87

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: Issued\_Patents\_AA:\*  
2: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/CTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	136	29.4	153	1	US-08-514-014-12
2	136	29.4	153	2	US-08-833-823-12
3	136	29.4	163	3	US-09-034-810-2
4	136	29.4	163	3	US-08-685-239-2
5	124	26.8	151	2	US-08-620-694A-8
6	124	26.8	151	3	US-09-034-810-6
7	124	26.8	151	3	US-09-022-255-8
8	124	26.8	151	3	US-09-022-696-8
9	124	26.8	151	3	US-08-685-239-6
10	124	26.8	151	3	US-09-022-253-8
11	124	26.8	151	3	US-09-022-260-8
12	119	25.7	150	3	US-09-034-810-4
13	119	25.7	150	3	US-08-685-239-4
14	119	25.7	158	2	US-08-620-694A-7
15	119	25.7	158	3	US-09-022-255-7
16	119	25.7	158	3	US-09-022-696-7
17	119	25.7	158	3	US-09-022-253-7
18	119	25.7	158	3	US-09-022-260-7
19	63	13.6	361	2	US-08-483-928A-9
20	57.5	12.4	1349	3	US-08-938-291A-6
21	57.5	12.4	362	2	US-08-737-045-13
22	56.5	12.2	455	1	US-08-476-008-3
23	56.5	12.2	455	1	US-08-306-063-3
24	56.5	12.2	455	1	US-08-833-485-3
25	56.5	12.2	455	4	PCR-US91-06148A-3
26	55	11.9	523	1	US-08-100-247-2
27	55	11.9	523	1	US-08-483-146A-2
28	55	11.9	523	1	US-08-233-513A-3

29	55	11.9	523	1	US-08-484-594A-2	Sequence 2, Appli
30	55	11.9	2439	3	US-09-335-409-7	Sequence 7, Appli
31	54.5	11.8	550	1	US-08-484-493-2	Sequence 2, Appli
32	54.5	11.8	550	1	US-08-484-494-2	Sequence 2, Appli
33	54.5	11.8	550	2	US-08-345-212-2	Sequence 2, Appli
34	54.5	11.8	550	3	US-09-249-003-2	Sequence 2, Appli
35	53	11.4	291	3	US-09-080-120A-7	Sequence 7, Appli
36	53	11.4	291	4	PCR-US95-08925-7	Sequence 7, Appli
37	52.5	11.3	349	3	US-08-630-172-6	Sequence 6, Appli
38	52.5	11.3	563	2	US-08-714-168-1	Sequence 1, Appli
39	52	11.2	1487	2	US-08-760-489-2	Sequence 2, Appli
40	52	11.2	1487	2	US-08-760-489-4	Sequence 4, Appli
41	51	11.0	95	2	US-08-722-349-3	Sequence 3, Appli
42	51	11.0	95	3	US-09-204-328-3	Sequence 3, Appli
43	51	11.0	741	1	US-08-277-231A-4	Sequence 3, Appli
44	51	11.0	741	1	US-08-473-750-7	Sequence 4, Appli
45	51	11.0	741	2	US-08-477-326-7	Sequence 7, Appli

## ALIGNMENTS

RESULT 1  
US-08-514-014-12  
Sequence 12, Application US/08514014  
Patent No. 5707829  
GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
APPLICANT: McCoy, John  
APPLICANT: Kelleher, Kerry  
APPLICANT: Carlin, McKeough  
TITLE OF INVENTION: DNA SEQUENCES AND SECRETED PROTEINS  
TITLE OF INVENTION: ENCODED THEREBY  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc. -- Legal Affairs  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/514,014  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Brown, Scott A.  
REGISTRATION NUMBER: 32,724  
REFERENCE/DOCKET NUMBER: G16000  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8224  
TELEFAX: (617) 876-5651  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 153 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-514-014-12

Query Match 29.4%; Score 136; DB 1; Length 153;  
Best Local Similarity 33.3%; Pred. No. 3.1e-11;  
Matches 33; Conservative 14; Mismatches 23; Indels 14; Gaps 3;  
QY 3 VDTEDRYPQKLAFAECLGCDIDARTGRTAALNSVRLQSLVLR--PCSRDGSGL 60  
DB 76 VYWDNRDPSEVVAQCRNLGCTINQ--GREDISMSVPIQETLVVRRKHGCS----- 128

```

QY      61  PTPGAFAFHTEFIHVPVGCCTVLP  84
          : | | : | | | | | | |
Db     129  ----VSFQLEKVLVTVGCTCTP  147

```

RESULT 2

```

US-08-833-823-12
Sequence 12, Application US/08833823
Patent No. 5965093
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John
APPLICANT: Kelleher, Kerry
APPLICANT: Carlini, McKeough
TITLE OF INVENTION: DNA SEQUENCES AND SECRETED PROTEINS
TITLE OF INVENTION: ENCODED THERBY
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech Institute, Inc. -- Legal Affairs
STREET: 87 Cambridgepark Drive
City: Cambridge
STATE: Massachusetts
COUNTRY: USA
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,823
FILING DATE: 10-APR-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/514,014
FILING DATE: 11-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Brown, Scott A.
REGISTRATION NUMBER: 32,724
REFERENCE/DOCKET NUMBER: 616000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8224
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 153 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-823-12

```

Query Match	29.4%	Score 136,	DB 2;	length 153;
Best Local Similarity	39.3%	Pred. NO. 3.1e-11;		
Matches 33;	Conservative 14;	Mismatches 23;	Indels 14;	Gaps 3;

```
QY      3 VDTEDRQRPKQLAAEECLRCGICDIARGRRETAALNSVALLQSLLVLRR--PCSRDSSGL 60
          :|::||::||::||::||::||::||::||::||::||::||::||::||
Db     76 VTMDPRNRPSRVQAQCNCNLCINAQ--GKEDIISNVSPIDQETLVARRKHGGCS----- 128

QY      61 PTPGAFAHTFEIHPVCCTVLP 84
          :|::||::||::||::||
Db    129 -----VSFQLEKVLVTGCCTVP 147
```

RESULT 3  
IIS-09-034-

Sequence 2, Application US/09034810  
Patent No. 6043344  
GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
APPLICANT: Kelleher, Kerry

```

1  APPLICANT: Carlin, McKeough
2  APPLICANT: Goldman, Samuel
3  APPLICANT: Pittman, Debra
4  APPLICANT: Mi, Sha
5  APPLICANT: Neben, Steven
6  APPLICANT: Giannotti, Joann
7  APPLICANT: Golden,Fleet, Margaret
8  TITLE OF INVENTION: Human CTLA-8 and Uses of CTLA-8-Related proteins
9  NUMBER OF SEQUENCES: 9
10 CORRESPONDENCE ADDRESS:
11 ADDRESSEE: Genetics Institute, Inc.
12 STREET: 87 Cambridgepark Drive
13 CITY: Cambridge
14 STATE: Massachusetts
15 COUNTRY: USA
16 ZIP: 02140
17
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: Floppy disk
20 COMPUTER: IBM PC compatible
21 OPERATING SYSTEM: PC-DOS/MS-DOS
22 SOFTWARE: PatentIn Release #1.0, Version #1.25
23 CURRENT APPLICATION DATA:
24 APPLICATION NUMBER: US/09/034,810
25 FILING DATE:
26 CLASSIFICATION:
27 PRIOR APPLICATION DATA:
28 APPLICATION NUMBER: 08/685,239
29 FILING DATE:
30 ATTORNEY/AGENT INFORMATION:
31 NAME: Brown, Scott A.
32 REGISTRATION NUMBER: 32,724
33 REFERENCE/DOCKET NUMBER: G15262
34 TELECOMMUNICATION INFORMATION:
35 TELEPHONE: (617) 498-8224
36 TELEFAX: (617) 876-5851
37
38 INFORMATION FOR SEQ ID NO: 2:
39 SEQUENCE CHARACTERISTICS:
40 LENGTH: 163 amino acids
41 TYPE: amino acid
42 TOPOLOGY: linear
43
44 MOLECULE TYPE: protein
45
46 US-09-034-810-2

```

Query Match	29.4%	Score 136;	DB 3;	Length 163;
Best Local Similarity	39.3%	Pred. No. 3.3e-11;		
Matches 33; Conservative	14;	Mismatches 23;	Indels 14;	Gaps 3;

```
QY      3 VDTEDERAPOLALAECCRCGCIADIRGRETAALNSVRLLOSLLVLR--PCSRDSSGL 60
        :|::||::||::||::||::||::||::||::||::||::||::||
Db     86 VTMDPNRRPSPVVAQAQCRCNIGCINAAQ-GKEDIISNVSPIDQETLVARRKHQGS----- 138
QY
Db    61 PTGGAFATHTFEIHVPVCCTVLP 84
       :|::||::||::||::||::||
   139 -----VSFQLEKVLVTGCCTVTP 157
```

RESULT

```

US-08-685-239-2
: Sequence 2, Application US/08685239
: Patent No. 6074849
:
: GENERAL INFORMATION:
:
: APPLICANT: Jacobs, Kenneth
: APPLICANT: Kelleher, Kerry
: APPLICANT: Carlin, McKeough
: APPLICANT: Goldman, Samuel
: APPLICANT: Pitman, Debra
: APPLICANT: Mi, Sha
: APPLICANT: Neben, Steven
: APPLICANT: Giannotti, Johnn
: APPLICANT: Golden/Fleet, Margaret
:
: TITLE OF INVENTION: Human CTLA-8 and Uses of CTLA-8-Related Proteins
:
: NUMBER OF SEQUENCES: 9

```







STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Herpesvirus Saimiri  
STRAIN: ORF13  
US-09-022-696-8

Query Match 26.8%; Score 124; DB 3; Length 151;  
Best Local Similarity 37.0%; Pred. No. 1.4e-09;  
Matches 30; Conservative 13; Mismatches 24; Indels 14; Gaps 3;

OY 6 DEDRYPQKLAFAECICRGCTDARTGRETALNSVRLQSLVLR--RCSRDGSGLP 63  
DB 77 DDDRPVSIVWEAKCRYGCVNA-DGNVDYHMSVPIQDEILVVRKGHCPCPN----- 127  
OY 64 GAFAPHEFIHVPGCTCVP 84  
DB 128 ---SFRLEKMLVVGCTCVP 145

RESULT 9  
US-08-685-239-6  
Sequence 6, Application US/08685239  
Patent No. 6074849  
GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
APPLICANT: Kelleher, Kerry  
APPLICANT: Carlin, McKeough  
APPLICANT: Goldman, Samuel  
APPLICANT: Pitman, Dedra  
APPLICANT: MI, Sha  
APPLICANT: Neben, Steven  
APPLICANT: Giannotti, Joann  
APPLICANT: Golden, Margaret  
TITLE OF INVENTION: Human CTLA-8 and Uses of CTLA-8-related Proteins  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/685,239  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Brown, Scott A.  
REGISTRATION NUMBER: 32,724  
REFERENCE/DOCKET NUMBER: G15262  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8224  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 151 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-685-239-6

Query Match 26.8%; Score 124; DB 3; Length 151;

Best Local Similarity 37.0%; Pred. No. 1.4e-09;  
Matches 30; Conservative 13; Mismatches 24; Indels 14; Gaps 3;

OY 6 DEDRYPQKLAFAECICRGCTDARTGRETALNSVRLQSLVLR--RCSRDGSGLP 63  
DB 77 DDDRPVSIVWEAKCRYGCVNA-DGNVDYHMSVPIQDEILVVRKGHCPCPN----- 127  
OY 64 GAFAPHEFIHVPGCTCVP 84  
DB 128 ---SFRLEKMLVVGCTCVP 145

RESULT 10  
US-09-022-253-8  
Sequence 8, Application US/09022253  
Patent No. 6096305  
GENERAL INFORMATION:  
APPLICANT: Yao, Zhengbin  
APPLICANT: Spriggs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 6096305e1 Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,253  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/620,694  
FILING DATE: 21-MARCH-1996  
APPLICATION NUMBER: USSN 08/538,765  
FILING DATE: 7 AUGUST 1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/410,535  
FILING DATE: 23 MARCH 1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 151 amino acids  
TYPE: amino acid  
STRANDEDNESS: No. 6096305 Relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Herpesvirus Saimiri  
STRAIN: ORF13  
US-09-022-253-8

Query Match 26.8%; Score 124; DB 3; Length 151;  
Best Local Similarity 37.0%; Pred. No. 1.4e-09;  
Matches 30; Conservative 13; Mismatches 24; Indels 14; Gaps 3;



Patent No. 6074849  
GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
APPLICANT: Kelleher, Kerry  
APPLICANT: Carlin, McKeough  
APPLICANT: Goldman, Samuel  
APPLICANT: Pitman, Debra  
APPLICANT: Ml, Sha  
APPLICANT: Neben, Steven  
APPLICANT: Giannotti, Joann  
APPLICANT: Golden, Fleet, Margaret  
TITLE OF INVENTION: Human CTLA-8 and uses of CTLA-8-related proteins  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/685,239  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Brown, Scott A.  
REGISTRATION NUMBER: 32,724  
REFERENCE/DOCKET NUMBER: G15262  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8224  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 150 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-685-239-4

Query Match 25.7%; Score 119; DB 3; Length 150;  
Best Local Similarity 38.0%; Pred. No. 6.7e-09;  
Matches 30; Conservative 11; Mismatches 24; Indels 14; Gaps 3;

QY 6 DEDRYPQKLAFAECRCIDARTGRTAALNSVRLQSLVLRPP--CSRDSGLPTP 63  
DB 76 DEDRIPSVIWEACRHRQCVNAE-GKLDHMSVLIQOEILVLRKPEKC----- 124  
QY 64 GAFAHTEFIHVPVGCIV 82  
DB 125 -PTEFVEKMLVGVGCTCV 142

RESULT 14  
US-08-620-694A-7  
Sequence 7, Application US/08620694A  
Patent No. 5869286  
GENERAL INFORMATION:  
APPLICANT: Yao, Zhengbin  
APPLICANT: Springs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 5869286el Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA

COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/620,694A  
FILING DATE: 21 MARCH 1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/538,765  
FILING DATE: 7 AUGUST 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/410,535  
FILING DATE: 23 MARCH 1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 158 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-620-694A-7

Query Match 25.7%; Score 119; DB 2; Length 158;  
Best Local Similarity 39.0%; Pred. No. 7.2e-09;  
Matches 30; Conservative 11; Mismatches 26; Indels 10; Gaps 2;

QY 6 DEDRYPQKLAFAECRCIDARTGRTAALNSVRLQSLVLRPP--CSRDSGLPTP 65  
DB 84 DEDRIPSVIWEACRHRQCVNAE-GKLDHMSVLIQOEILVLRKPEPS-----CP 133  
QY 66 FAHTEFIHVPVGCIV 82  
DB 134 FTEFVEKMLVGVGCTCV 150

RESULT 15  
US-09-022-255-7  
Sequence 7, Application US/09022255  
Patent No. 6072033  
GENERAL INFORMATION:  
APPLICANT: Yao, Zhengbin  
APPLICANT: Springs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 6072033el Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,255  
FILING DATE:

```

;
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/620,694
; FILING DATE: 21 MARCH 1996
; APPLICATION NUMBER: USSN 08/538,765
; FILING DATE: 7 AUGUST 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/410,535
; FILING DATE: 23 MARCH 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Perkins, Patricia Anne
; REGISTRATION NUMBER: 34,695
; REFERENCE/DOCKET NUMBER: 2617-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)587-0430
; TELEFAX: (206)
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-09-022-255-7

```

```

Query Match      25.7%; Score 119; DB 3; Length 158;
Best Local Similarity 39.0%; Pred. No. 7.2e-09;
Matches 30; Conservative 11; Mismatches 26; Indels 10; Gaps 2;

```

```

QY 6 DEDRYPOKLAFAECLRCGCIARTGRTAALNSVRLQSLVLRRRPCSRDGSGLPTPGA 65
   | | | | | : | : | : | : | : | : | : | : | : | : | : | : |
Db 84 DPDRYPVYWEAQCRHRCVNAE-GKLDHMHNSVLIQOEILVLKREPES-----CP 133

QY 66 FAFHTEFIHVPVGCVCV 82
   | | | | | : | : | : | : | : | : | : | : | : | : | : | : |
Db 134 FTFRYEKMLVGVGCTCV 150

```

Search completed: February 16, 2001, 12:50:10  
 Job time: 130 sec

Gencore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2001, 20:41:30 ; Search time 2862.16 Seconds  
(without alignments)  
2936.015 Million cell updates/sec

Title: US-09-320-713-3  
Perfect score: 1642  
Sequence: 1 ggaattcgccagcagctcgtl.....tgaaaaaaaaaaaaaaaaaa 1642

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 1118133 seqs, 255875100 residues

Total number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%  
Listing first 45 summaries

Database :

GenBank: 1: gb\_ba1:\*  
2: gb\_ba2:\*  
3: gb\_cm:\*  
4: gb\_ov:\*  
5: gb\_ph:\*  
6: gb\_pl1:\*  
7: gb\_pl2:\*  
8: gb\_pr1:\*  
9: gb\_pr2:\*  
10: gb\_pr3:\*  
11: em\_fun:\*  
12: em\_hum1:\*  
13: em\_hum2:\*  
14: em\_in:\*  
15: em\_cm:\*  
16: em\_or:\*  
17: em\_ov:\*  
18: em\_pal:\*  
19: em\_ph:\*  
20: em\_pl:\*  
21: em\_ro:\*  
22: em\_sts:\*  
23: em\_sy:\*  
24: em\_un:\*  
25: em\_vl:\*  
26: gb\_hcg1:\*  
27: gb\_hcg2:\*  
28: gb\_in1:\*  
29: gb\_in2:\*  
30: em\_ba1:\*  
31: em\_ba2:\*  
32: em\_hum3:\*  
33: em\_hum4:\*  
34: gb\_pr4:\*  
35: gb\_hcg3:\*  
36: gb\_hcg4:\*  
37: gb\_hcg5:\*  
38: gb\_hcg6:\*  
39: gb\_hcg7:\*  
40: em\_hcg1:\*  
41: em\_hcg2:\*  
42: em\_hcg3:\*  
43: em\_hum5:\*

44: gb\_pl3:\*  
45: gb\_pr5:\*  
46: gb\_hcg8:\*  
47: gb\_hcg9:\*  
48: gb\_hcg10:\*  
49: gb\_hcg11:\*  
50: gb\_hcg12:\*  
51: gb\_hcg13:\*  
52: gb\_hcg14:\*  
53: gb\_in3:\*  
54: gb\_hcg15:\*  
55: gb\_hcg16:\*  
56: gb\_hcg17:\*  
57: em\_hcg4:\*  
58: em\_hcg5:\*  
59: em\_hcg6:\*  
60: em\_hcg7:\*  
61: em\_hum6:\*  
62: gb\_hcg18:\*  
63: gb\_hcg19:\*  
64: gb\_ba3:\*  
65: em\_hcg8:\*  
66: em\_hcg9:\*  
67: em\_hcg10:\*  
68: gb\_pr6:\*  
69: gb\_pr7:\*  
70: gb\_hcg20:\*  
71: gb\_hcg21:\*  
72: gb\_hcg22:\*  
73: gb\_hcg23:\*  
74: gb\_ro:\*  
75: gb\_sts1:\*  
76: gb\_sts2:\*  
77: gb\_sy:\*  
78: gb\_un:\*  
79: gb\_vl1:\*  
80: gb\_vl2:\*  
81: gb\_pal1:\*  
82: gb\_pal2:\*  
83: em\_hcg0:\*  
84: gb\_hcg24:\*  
85: gb\_pr8:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	1472.2	89.7	140800	78	AC018392	AC018392 Homo sapi
c 2	1455.2	88.6	142376	80	AL161772	AL161772 Homo sapi
c 3	624.4	38.0	629	85	HUMNOR1B	L23207 Human chrom
c 4	294.4	17.9	286	85	HUMCH13C	L30108 Homo sapien
c 5	252.8	15.4	323	85	HUMNOR1A	L23206 Human chrom
c 6	245.2	14.9	250	75	G31083	G31083 human STS S
c 7	77	4.7	162005	71	AL356002	AL356002 Homo sapi
c 8	73.2	4.5	687	34	AF152098	AF152098 Homo sapi
c 9	73.2	4.5	688	34	AF184969	AF184969 Homo sapi
c 10	73.2	4.5	711	34	AF212311	AF212311 Homo sapi
c 11	72.8	4.4	42655	64	SC7H2	AL109732 Streptomy
c 12	65	4.0	1429	7	D76415	D76415 Rice seedl
c 13	65	4.0	2588	6	AF099203	AF099203 Oryza sat
c 14	63.6	3.9	3578	1	AF005277	AF005277 Cellulomo
c 15	63.6	3.9	180630	49	AC025783	AC025783 Oryza sat
c 16	63.6	3.9	224517	56	AC074335	AC074335 Mus muscu
c 17	63.4	3.9	332976	63	AC084055	AC084055 Mus muscu
c 18	62.6	3.8	144500	35	AC010231	AC010231 Homo sapi
c 19	61.4	3.7	4615	64	STMWH1B12X	L22864 Streptomyce
c 20	61.4	3.7	38390	64	SC2H12	AL359215 Streptomy
c 21	61	3.7	300695	62	AC079431	AC079431 Mus muscu

```

22 60.8 3.7 143209 7 AP001550
23 60.6 3.7 8058 SCE56
24 60.2 3.7 33810 64 SCE5
25 60.2 3.7 77457 1 AF210249
26 60 3.7 2544 6 AB004819
27 59.6 3.6 2381 34 AK025816
28 59.6 3.6 2953 34 AK025470
29 59.4 3.6 692 74 AF184970
30 59.4 3.6 14634 64 SC8B7
31 58.8 3.6 78220 47 AC023212
32 58.6 3.6 30991 64 SCC53
33 58.4 3.6 150074 37 AC016007
34 58.2 3.5 38426 64 SCF55
35 58.2 3.5 107109 69 HSDJ781B1
36 57.8 3.5 165943 47 AC022205
37 57.4 3.5 190014 7 AP000836
38 57 3.5 200580 63 AC084064
39 56.8 3.5 1047 34 AF152099
40 56.8 3.5 22449 64 SPSNBCDE
41 56.8 3.5 22449 64 SPSNBCGEN
42 56.8 3.5 68848 2 MTW043
43 56.8 3.5 189579 48 AC023885
44 56.8 3.5 245035 50 AC026782
45 56.6 3.4 154868 35 AC011316

```

## ALIGNMENTS

```

RESULT 1
AC018392/c AC018392 140800 bp DNA HTG 09-SEP-2000
DEFINITION Homo sapiens clone RP11-12K9, WORKING DRAFT SEQUENCE, 14 unordered
LOCUS AC018392.4 GI:10045345
ACCESSION AC018392
VERSION AC018392.4
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens

```

```

REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
TITLE 1 (bases 1 to 140800)
JOURNAL Homo sapiens, clone RP11-12K9
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 140800)

```

```

TITLE Direct Submission
JOURNAL Submitted (10-DEC-1999) Whitehead Institute/MIT Center for Genome
COMMENT Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 9, 2000 this sequence version replaced gi:6649406.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

```

```

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

```

```

----- Project Information
Center project name: L3288
Center clone name: L2_K_9
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 133050 bases at least Q40
Consensus quality: 136489 bases at least Q30
Consensus quality: 137937 bases at least Q20
Insert size: 139000; agarose-fp
Insert size: 139500; sum-of-ctnigs
Quality coverage: 5.3 in Q20 bases; agarose-fp
Quality coverage: 5.3 in Q20 bases; sum-of-ctnigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1. 4056: contig of 4056 bp in length.
* 4057 4156: gap of 100 bp
* 4157 5294: contig of 1138 bp in length
* 5295 5394: gap of 100 bp
* 5395 6816: contig of 1422 bp in length
* 6817 6916: gap of 100 bp
* 6917 57039: contig of 50123 bp in length
* 57040 57139: gap of 100 bp
* 57140 60223: contig of 3084 bp in length
* 60224 60323: gap of 100 bp
* 60324 63674: contig of 3351 bp in length
* 63675 63774: gap of 100 bp
* 63775 68200: contig of 4426 bp in length
* 68201 68300: gap of 100 bp
* 68301 76558: contig of 8258 bp in length
* 76559 76658: gap of 100 bp
* 76659 84493: contig of 7835 bp in length
* 84494 84593: gap of 100 bp
* 84594 94397: contig of 9804 bp in length
* 94398 94497: gap of 100 bp
* 94498 104204: contig of 9707 bp in length
* 104205 104304: gap of 100 bp
* 104305 118656: contig of 14352 bp in length
* 118657 118756: gap of 100 bp
* 118757 139302: contig of 20546 bp in length
* 139303 139402: gap of 100 bp
* 139403 140800: contig of 1398 bp in length.
Location/Qualifiers
1. 140800
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-12K9"
/clone_11b="RPCT-11 Human Male BAC"
1. 4056
/note="assembly-fragment"
clone_end:SP6
vector_side:left"
misc_feature
4157..5294
/note="assembly-fragment"
5395..6816
/note="assembly-fragment"
6917..57039
/note="assembly-fragment"
57140..60223
/note="assembly-fragment"
60324..63674
/note="assembly-fragment"
63775..68200
/note="assembly-fragment"
68301..76558

```







Db 130960 TCCCCACGCTGCTGCGCCGACACCCGCCCTGCGCGCGCGCTTCCGCTACACCG 130901  
 QY 349 aggcctacgtacacatcccgctggagctgacatgctcccgagccgagagagagagagag 408  
 Db 130900 AGGCTTACGTACCAATCCCTGGGCTGACCTGCTGCCGAGCCGAGAGAGAGAGAGAG 130841  
 QY 409 aacagatcaactcagacatcgacacaacagagcgcaagctcctgctgagcccaacagag 468  
 Db 130840 ACAGATCAACTCAGCATGCAACAAGAGGCGCAACCTCTGCTGGGCCCAACGACG 130781  
 QY 469 cgcgcgctgccccctgagccgctgccccgagagagctcccgccgagccgagccgag 528  
 Db 130780 CGCCCGCTGGCCCTGAGAGCGCGCTGCTCCCGGAGGCTCTCCCGCGCCCATCCGAG 130721  
 QY 529 gcgcaccaagctggagccgctggagagagctgctggagacatcctgaaagagctgaccca 588  
 Db 130720 GCGCCCAAGCTGGAGCCCGCTGGAGAGGCTGCTGCGGCACTCTGAAGAAGAGTCCACCGA 130661  
 QY 589 gcaaaccaagctgacgagacacagcgccgcttccacagagactcgtaaagcagcttcat 648  
 Db 130660 GCAAAACCAAGTGCAGGACACACAGCGCGCTTCCATGAGAGACTCTGAACGACTTCAT 130601  
 QY 649 ctgacacagggacatccctgctgctttagctacaagacagcagcgtgctgagagctga 708  
 Db 130600 CTGACACGCGCATCCCTGCTGCTTTAGCTACAAAGCAACGACGCTGGAAGCTGA 130541  
 QY 709 tgggaaagacccgagcagggcagctcgtgtgfcgagcccgacagagaggttggaaaagt 768  
 Db 130540 TGGGAAGACGACCCGCGACGGGCTATCTGTGCGCCGCGCAATGAGAGGTTTGGAAAGTT 130481  
 QY 769 caagagagctccctgagagagcctcagatcgtcgtgctgagagcgagagcgctgactcac 828  
 Db 130480 CACGAGAGCTCCCTGAGAGAGCTCTGCAAGATGGCGTGGCGGAGGAGGAGGCTGACTCAC 130421  
 QY 829 cgcctggtgctgctgcaagagatagagagcagatgcttcttaagcaatcctaaataa 888  
 Db 130420 CGCTGGGCTGCTGCAAGAGATGAGGAGCATATGCTTTTAAAGCAATCTAAAGATTA 130361  
 QY 889 taataagatatgagactatatacctactttaaatacaactgcttgaatagagagcagag 948  
 Db 130360 TAAATAGATATAGCAGATATATACCTACTTTAAATCAACGTTTGAATAGAGGCAAG 130301  
 QY 949 ctatttataatcaaatgagagcactcgtctacatcttcttaacatataacatcgtt 1008  
 Db 130300 CTATTTATATATCAATGAGAGACTCTGTTACATTTCTTAAATATTAACATCGTT 130241  
 QY 1009 tttaactctctgctgagaaattttaaagcaaatggaatcctgagataatttgta 1068  
 Db 130240 TTTTACTTCTCTGATGAAATTTTAAAGCATATTTGAATCTTGGAATTTTGTGTA 130181  
 QY 1069 gctggtacactgctggcgtggtctctgaattcagccgtctacacgagctgagcgtatgaa 1128  
 Db 130180 GCTGATACGCTGCGGCTGCTGTAATTCACCTGCTCACCGATGCTGACATGATGAA 130121  
 QY 1129 atggacacgtctcatctgacacacactctcctcctcagaaagcttcaacgagcctcag 1188  
 Db 130120 ATGGAACAGTCTCATGAGACCACTCTTCTTCACATGAGAGCTTTCACGCGCTCCAG 130061  
 QY 1189 tggaccaaaggatgacacagcgctgcatgccccagggcagctaaagattccaaaga 1248  
 Db 130060 TGGACCAAAAGGATGACACAGCGGCTGCAATGCCCCAGGCGCACTAAAGATTTCCAAGA 130001  
 QY 1249 tctagatttggttttagtcaatgaataaagctcaaatctgcaaatcttcc 1308  
 Db 130000 TCTAGATTTGGTTTGTAGTCAATGAATATTAACAGTCTCAAACTCCGACATTTTTC 129941  
 QY 1309 ccccttggaaagcacttggggccaattgltgltaaagagctggtgagataaagagtgaa 1368  
 Db 129940 CCCTTTTGAAGCACTGGGGCCAAATTTGTGTTAAAGGAGTGAGATTAAGAGTGGA 129881  
 QY 1369 cgtgacatcttgcagctgtgcagaaagatcacaagcaggtattggttagtgaagagc 1428  
 Db 129880 CGTGACATCTTTGCGAGTTGTGCAAGAAATCCAAAGCAGGTATTGGCTTAAAGGCG 129821

QY 1429 tttagatcagcgtcgtgaatagagacaagtggtgcaacgttagatctgcagatcaat 1488  
 Db 129820 TTTAGATCAGCGCTGAATATGAGACAAAGTGGCGCACGTTTAGCATCTGCAAGATCAAT 129761  
 QY 1489 ctgagagctctgcttctgcattctgcacagagagcaggtcctgacttcttcttga 1548  
 Db 129760 CTGAAGGCTCTGTCTTCTGCAATTTGCCACGAGAGCTAGGCTCTTGATCTTTTCTTGA 129701  
 QY 1549 ttgaagctgctctgaaacaattatgttgaagttgtagtcttctttaaactcat 1608  
 Db 129700 TTGAAGCTGTCTTTCATCTTCATATATATGATAAAGTAAAGTAACTTTTAAATCAT 129641  
 QY 1609 taaagagcgtctgtgaaaaaaa 1633  
 Db 129640 TAAAGAGGCTTGTCTGAAGATACA 129616

RESULT 3  
 HUMNOTIB/c  
 LOCUS Human chromosome-specific mRNA.  
 DEFINITION Human chromosome-specific mRNA.  
 ACCESSION I23207  
 VERSION I23207.1 GI:434048  
 KEYWORDS  
 SOURCE Homo sapiens female 3 month post natal entire brain cdna to mRNA.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 AUTHORS Bonaldo, M., Soares, M.-B. and Warburton, D.  
 TITLE Selection of chromosome-specific cDNAs and their corresponding  
 genomic clones  
 JOURNAL Unpublished (1993)  
 COMMENT Chromosome 13q11.  
 FEATURES  
 source location/Qualifiers  
 1. 629  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /dev\_stage="3 month post natal"  
 /sex="female"  
 /tissue\_type="entire brain"

BASE COUNT 189 a 149 c 128 g 163 t  
 ORIGIN

Query Match 38.0%; Score 624.4; DB 85; Length 629;  
 Best Local Similarity 99.8%; Pred. No. 5, 6e-87;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1001 acatcgttcttactctctcgtgtagaattttaaagcaaatggaatccttgaata 1060  
 Db 629 ACATCGTTTTTACTCTCTGAGAAATTTGTTAAAGCAATTTGAATCTTGATTA 570  
 QY 1061 attttagctgtagaactctgctggtgctcgtgaatgaagcctgacagagtgta 1120  
 Db 569 ATTTTGAAGTGTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 510  
 QY 1121 ctgataaatgagacgtctcatctgacacacacttctctcactgagaggttctcaag 1180  
 Db 509 CTGATGAATAGGACACGTCATCTGACCACTCTTCCACTGAAGGTCTTACAGCG 450  
 QY 1181 ccttcaggtgagcaaaaggatgacagcgctgcatgccccagggcagctaaagat 1240  
 Db 449 CCTCAGGTGAGCAAAAGGATGACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 390  
 QY 1241 tccaaagatcagagttttagtcatgataaacaagctcaaatctgcaacacgacaa 1300  
 Db 389 TCCAAAATCTCAGATTTGTTTATGATGAATATTAACAGTCTCAAACTGCAACA 330  
 QY 1301 ttcttcccttcttgaagcactggtggccaatttggttlaagagtggtgagataag 1360  
 Db 329 TTTTTCCTCCCTTTTGAAGCACTGGGCGCAATTTGTGTTAAGAGTGTGATGAAG 270



Stanford University School of Medicine  
Department of Genetics, M-344, Stanford, CA 94305, USA  
Tel: 415/7259687  
Fax: 415/7259689  
Email: myers@shgc.stanford.edu

Primer A: CAAGGACCTACCTCTGCGG  
Primer B: GAAGATCCACGACGATATG  
STS size: 143

PCR Profile:

Initial incubation: 94 degrees C for 90 seconds

Denaturation: 94 degrees C for 15 seconds

Annealing: 62 degrees C for 23 seconds

Polymerization: 72 degrees C for 30 seconds

PCR Cycles: 30

Thermal Cycler: Perkin Elmer 9600

Protocol:

Template: 25 ng

Primer: each 1 uM

dNTPs: each 200 uM

Taq Polymerase: 0.05 units/uL

Total Vol: 10 uL

Buffer:

MgCl<sub>2</sub>: 2.5 mM

KCl: 50 mM

Tris-HCl: 20 mM

pH: 8.3

FEATURES  
Prepared with primer pairs derived from H29012--Merck/UniEST.  
location/Qualifiers

1..250

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/map="13"

STS

primer\_bind 108..250

primer\_bind 108..127

complement(229..250)

BASE COUNT 77 a 58 c 38 g 77 t

ORIGIN

Query Match

14.9% Score: 245.2; DB: 75; Length: 250;

Best Local Similarity 98.8%; Pred. No. 1.5e-28;

Matches 247; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1392 gaagaatccaagcagatctgcttaagttgtaagccttagatccaagcctgaatcgaag 1451

Db 250 GAAGAATCCAAGCAGATCTGCTTAAGTTGTAAGCCTTAGATCCAAGCCTGAATATGAG 191

QY 1452 gacaaagtggcgacgcttaagcattcagagatcaatctggaagctctctgttcgatt 1511

Db 190 GACAAAGTGGCGCACGCTTACATCTGCAGATCAATCTGGAGCCTTCTGTCATT 131

QY 1512 ctgcacagagaagcagctctgcttcttcttagatgaaagctcgtctctgaacaca 1571

Db 130 CTGCACAGAGAGCTAGCTCTGATCTTCTTAAAGTCTGCTCTGAACACA 71

QY 1572 attattgtaaaagttagtctctcttaaatcattaaagagcctgctgaacaaa 1631

Db 70 ATTATTTGTAAGTGTAGTGTCTTTTAAATCAATTAAGAGGCTTCTGTAAGATA 11

QY 1632 aaaaaa 1641

Db 10 AAAAAAAAAA 1

RESULT 7

AL356002/c 162005 bp DNA

LOCUS Homo sapiens chromosome 1 clone RP11-324K19, \*\*\* SEQUENCING IN

DEFINITION PROGRESS \*\*\* 18 unordered pieces.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AL356002  
AL356002.4 GI:9797411  
HTG, HTGS\_PHASE1, HTGS\_DRAFT.  
human.  
Homo sapiens

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Submitted (09-AUG-2000) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
On Aug 12, 2000 this sequence version replaced gi:9713703.  
Genome Center  
Center: Sanger Centre  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
----- Project Information  
Center project name: DA324K19  
----- Summary Statistics  
Assembly program: XGAP4; Version 4.5  
Sequencing vector: plasmid; L08752; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Consensus quality: 150712 bases at least Q40  
Consensus quality: 155209 bases at least Q30  
Consensus quality: 157665 bases at least Q20  
Insert size: 160305; sum-of-ctrls  
Insert size: 170079; 1.7% error; agarose-fp  
Quality coverage: 3.29x in Q20 bases; sum-of-ctrls Quality  
coverage: 3.10x in Q20 bases; agarose-fp

COMMENT

NOTE: This is a 'working draft' sequence. It currently  
consists of 18 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1 4977: contig of 4977 bp in length  
\* 4978 5077: gap of 100 bp  
\* 5078 14284: contig of 9207 bp in length  
\* 14285 14384: gap of 100 bp  
\* 14385 20588: contig of 6204 bp in length  
\* 20589 20688: gap of 100 bp  
\* 20689 27191: contig of 6503 bp in length  
\* 27192 27291: gap of 100 bp  
\* 27292 39205: contig of 11914 bp in length  
\* 39206 39305: gap of 100 bp  
\* 39306 43223: contig of 3918 bp in length  
\* 43224 43323: gap of 100 bp  
\* 43324 46284: contig of 2961 bp in length  
\* 46285 46384: gap of 100 bp  
\* 46385 73430: contig of 27046 bp in length  
\* 73431 73530: gap of 100 bp  
\* 73531 104566: contig of 31036 bp in length  
\* 104567 104666: gap of 100 bp  
\* 104667 109641: contig of 4975 bp in length  
\* 109642 109741: gap of 100 bp  
\* 109742 113089: contig of 3348 bp in length  
\* 113090 113189: gap of 100 bp  
\* 113190 132251: contig of 19062 bp in length  
\* 132252 132351: gap of 100 bp  
\* 132352 138017: contig of 5666 bp in length  
\* 138018 138117: gap of 100 bp  
\* 138118 141454: contig of 3337 bp in length  
\* 141455 141554: gap of 100 bp  
\* 141555 145269: contig of 3715 bp in length  
\* 145270 145369: gap of 100 bp  
\* 145370 150364: contig of 4995 bp in length  
\* 150365 150464: gap of 100 bp



BASE COUNT	142 a	216 c	209 g	120 t
ORIGIN	GPCRORAMETIAVGCYCF"			





```

stem_loop
/ gene="SC7H2.02"
/ note="Pfam match to entry PF00583 Acetyltransferase,
Acetyltransferase (GNAT) family, score 43.70, E-value
4.2e-09."
820..900
/ note="Inverted repeat (Score 57: 23/26 ( 88%) matches, 0
gaps) with potential 28 base loop."
complement(907..1878)
/ gene="SC7H2.03c"
complement(907..1878)
/ gene="SC7H2.03c"
/ note="SC7H2.03c, probable aminopeptidase, len: 323aa;
similar to many proline aminopeptidases eg. SW:PIP_XMCI
proline iminopeptidase from Xanthomonas campestris pv.
citri (313 aa) fasta scores: opt: 848, z-score: 979.3,
E(): 0, (44.4% identity in 320 aa overlap). Contains Pfam
match to entry PF00561 abhydrolase, alpha/beta hydrolase
fold."
/ codon_start=1
/ trans1_table=11
/ label="SC7H2.03c"
/ product="putative aminopeptidase"
/ protein_id="CAB52045.1"
/ db_xref="GI:5689882"
/ translation="MSLYPEIEPYDHGMLDVGDNHVMYETCGNPGKPAVYLHGSG
SRASPLRFPDPAAYRIVLDOGRGRLSPASAPDPTMSVTATHLMADELRLRFLVH
LGIERLIWNGVSMGSLGLRYAQTDFGVTELVITVAATGSAEVALIRGLANTPE
AHERFLAELPPDARDGNLPAAYRRLSPDPAKREARAWTDWEATIPAPGSVAR
YDDPDRMGFAFRTYHYWGNDFLADGNDGVYIRAHLLKGIPTGLVQSGDFGNIL
GIYWRLHAMPSDILVIDEAGHDAGTGDDEALLATDKYARNGTAE"
complement(928..1701)
/ gene="SC7H2.03c"
/ note="Pfam match to entry PF00561 abhydrolase, alpha/beta
hydrolase fold, score 74.20, E-value 2.7e-18."
complement(1933..2646)
/ gene="SC7H2.04c"
complement(1933..2646)
/ gene="SC7H2.04c"
/ note="SC7H2.04c, unknown, len: 237aa; similar to
SW:Y0BN_MTCU hypothetical protein from Mycobacterium
tuberculosis (350 aa) fasta scores: opt: 870, z-score:
990.4, E(): 0, (57.1% identity in 231 aa overlap)."
/ codon_start=1
/ trans1_table=11
/ label="SC7H2.04c"
/ product="hypothetical protein"
/ protein_id="CAB52046.1"
/ db_xref="GI:5689883"
/ translation="MDTGGISVLDRIIEGCRACPLVEMREEVARTKRAAFADWTYMG
RPVPGRPDPDARLLIYGLAPAHGNGNRMTGDSGVITQALDVGLSQPTAVR
YDDGLELYGVRTSPVHCAPPAKPKPTPAERDTCRSLVDELGLRTLLRAVYVLGAFG
WDAALPFAAGAGTVPKPRPAFHGTOTLDAADGPDHLFGCFHVSORNTFTGRLLP
EMLRDVLRTAETAGLPAAR"
2763..3323
/ gene="SC7H2.05"
2763..3323
/ gene="SC7H2.05"
/ note="SC7H2.05, unknown, len: 186aa; similar to
SW:YRPH_ECOLI hypothetical protein from Escherichia coli
(133 aa) fasta scores: opt: 228, z-score: 249.5, E():
1.4e-06, (35.2% identity in 122 aa overlap)."
/ codon_start=1
/ trans1_table=11
/ label="SC7H2.05"
/ product="hypothetical protein"
/ protein_id="CAB52047.1"
/ db_xref="GI:5689884"
/ translation="MASEHDEGSGPVAAPROCAAAAGTPDKTATPDPKAAATA
AARAGQHSESVAVDSMIVAVRLIKTRISGATACRCHRVANERKAPVSLRQGEV
KLKQEGREVRVAVKRLIRKRVGAPVAVQCYVDSPPPPREAVAPAGINDRGAGRPTK
RDRREMDRLRGLEGLSSRRDAGTRP"
complement(3375..3881)
/ gene="SC7H2.06c"

```

```

CDS
complement(3375..3881)
/ gene="SC7H2.06c"
/ note="SC7H2.06c, possible integral membrane protein, len:
186aa; contains possible membrane spanning hydrophobic
regions."
/ codon_start=1
/ trans1_table=11
/ label="SC7H2.06c"
/ product="putative integral membrane protein"
/ protein_id="CAB52048.1"

Query Match 4.4%; Score 72.8; DB 64; Length 42655;
Best Local Similarity 46.7%; Pred. No. 0.019;
Matches 230; Conservative 0; Mismatches 262; Indels 0; Gaps 0;

QY 66 gagcagcgcgcaacgcaagcgtgcccggcaggggagcccgccgacgcgttcg 125
|||||
Db 37276 GAGCTGCGGAGACGACCCGACCGCACCGGCGAGGAGGCGTCCGCGCGCGCG

QY 126 cgcggcccaacccctgcgcagcgtgtgccttgccctacgaatctctagaccggcg 185
|||||
Db 37216 CTGCGGAGACCGGAGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG

QY 186 aggtaccaccaagctgcctggaagctactgcctgctgcgggggctgacctgacgggctg 245
|||||
Db 37156 TCCAACTACCTGTTACACCGCCACCGCCGAGGCGCTTGTGGACAGCCGAAAGCGGAGCTG

QY 246 ttcggcagaggaagcgtgcgtcttcgcagcgcccgctctacatgcccacgtgctcgt 305
|||||
Db 37096 GTCCGGAGCTACGACGACCGGCTACTACCCGAGGCGGCTGCGGCGGCGGCGGCGG

QY 306 cgcgcgacccccgcctgcctgcggcgcgccgtctcgtctacacgagcgctacgctaccatc 365
|||||
Db 37036 CCGCGCATGGCCCAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG

QY 366 ccggtggcgtcacctcgtgcccgagcgaggaagcagcagcagcagcagcagcagcagc 425
|||||
Db 36976 ACCCTGGCGCTGGCGCGGAGTGCCTGCGGACCGCGCATCCGATCCCGCGCGCGCGCG

QY 426 atgcacaagaagcgcaagctcgtcgtgcccccaagcagcgccgctgagccctga 485
|||||
Db 36916 AAGCTCACACGACGCTGACGACGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG

QY 486 ggcgcgtcctgcgcgcggagaggtctcccgcccgcatcccgagcgcccaagctgagcgc 545
|||||
Db 36856 GACTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG

QY 546 gctctgagggcct 557
|||||
Db 36796 AGCCGCGCGGAT 36785

RESULT 12
LOCUS D76415 1429 bp mRNA PLN 01-FEB-2000
DEFINITION Rice seedling mRNA for cysteine proteinase, complete cds.
ACCESSION D76415
VERSION D76415.1 GI:1514952
KEYWORDS cysteine proteinase.
SOURCE Oryza sativa (strain:cv. Koshihikari) seedling aleurone layer and
endosperm cDNA to mRNA, clone:PREPI.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Poales;
Poaceae; Oryza.
1 (bases 1 to 1429)
AUTHORS Kato, H. and Minamikawa, T.
TITLE Identification and characterization of a rice cysteine
endopeptidase that digests glutelin
JOURNAL Eur. J. Biochem. 239 (2), 310-316 (1996)
MEDLINE 96314485
REFERENCE 2 (bases 1 to 1429)
AUTHORS Kato, H.

```









117: em\_estp16:\*  
 118: em\_estp17:\*  
 119: em\_estp18:\*  
 120: em\_estp19:\*  
 121: em\_estp20:\*  
 122: em\_estp21:\*  
 123: em\_estp22:\*  
 124: em\_estp23:\*  
 125: em\_estp24:\*  
 126: em\_estp25:\*  
 127: em\_estp26:\*  
 128: em\_estp27:\*  
 129: em\_estp28:\*  
 130: em\_estp29:\*  
 131: em\_estp30:\*  
 132: em\_estp31:\*  
 133: em\_estp32:\*  
 134: em\_estp33:\*  
 135: em\_estp34:\*  
 136: em\_estp35:\*  
 137: em\_estp36:\*  
 138: em\_estp37:\*  
 139: em\_estp38:\*  
 140: em\_estp39:\*  
 141: em\_estp40:\*  
 142: em\_estp41:\*  
 143: em\_estp42:\*  
 144: em\_estp43:\*  
 145: em\_estp44:\*  
 146: em\_estp45:\*  
 147: em\_estp46:\*  
 148: em\_estp47:\*  
 149: em\_estp48:\*  
 150: em\_estp49:\*  
 151: em\_estp50:\*  
 152: em\_estp51:\*  
 153: em\_estp52:\*  
 154: em\_estp53:\*  
 155: em\_estp54:\*  
 156: em\_estp55:\*  
 157: em\_estp56:\*  
 158: em\_estp57:\*  
 159: em\_estp58:\*  
 160: em\_estp59:\*  
 161: em\_estp60:\*  
 162: em\_estp61:\*  
 163: em\_estp62:\*  
 164: em\_estp63:\*  
 165: em\_estp64:\*  
 166: em\_estp65:\*  
 167: em\_estp66:\*  
 168: em\_estp67:\*  
 169: em\_estp68:\*  
 170: em\_estp69:\*  
 171: em\_estp70:\*  
 172: em\_estp71:\*  
 173: em\_estp72:\*  
 174: em\_estp73:\*  
 175: em\_estp74:\*  
 176: em\_estp75:\*  
 177: em\_estp76:\*  
 178: em\_estp77:\*  
 179: em\_estp78:\*  
 180: em\_estp79:\*  
 181: em\_estp80:\*  
 182: em\_estp81:\*  
 183: em\_estp82:\*  
 184: em\_estp83:\*  
 185: em\_estp84:\*  
 186: em\_estp85:\*  
 187: em\_estp86:\*  
 188: em\_estp87:\*  
 189: em\_estp88:\*

190: gb\_gss25:\*  
 191: gb\_gss26:\*  
 192: gb\_gss27:\*  
 193: gb\_gss28:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	579.6	35.3	586	136	BE856748	BE856748 7f68b12.x
2	545.4	33.2	579	97	AW956791	AW956791 EST368661
3	527	32.1	529	26	AI870335	AI870335 w172f12.x
4	517.2	31.5	532	108	BE465289	BE465289 hw14c08.x
5	480.6	29.3	503	38	AW015888	AW015888 UI-H-B10-
6	467	28.4	476	25	AI817213	AI817213 wg62h12.x
7	456.2	27.8	461	11	AA778133	AA778133 z146h06.s
8	452	27.5	487	21	AI492261	AI492261 t159b03.x
9	448.8	27.3	466	143	NS2801	NS2801 v218g05.s1
10	448.2	27.3	453	93	AW614686	AW614686 hb31b08.x
11	442.4	26.9	446	22	AI569688	AI569688 tn04e05.x
12	439.8	26.8	443	17	AI199159	AI199159 q146h07.x
13	421.8	25.7	475	143	NS2873	NS2873 zb71d08.s1
14	420.2	25.6	425	19	AI371173	AI371173 ta09c11.x
15	417	25.4	419	16	AI085287	AI085287 qf18f01.x
16	413.8	25.2	425	16	AI146504	AI146504 qb84h09.x
17	411.8	25.1	469	141	H29012	H29012 ym31h04.s1
18	410.2	25.0	425	24	AI740658	AI740658 wg07a09.x
19	408.6	24.9	427	87	AW236638	AW236638 xm48h05.x
20	407.6	24.8	426	17	AI206346	AI206346 q121h05.x
21	405.6	24.7	414	141	P36829	P36829 HSPD34729.H
22	398.4	24.3	441	18	AI275406	AI275406 q163c10.x
23	396.4	24.1	406	19	AI339872	AI339872 q43d12.x
24	395	24.1	465	141	H30762	H30762 y079b08.r1
25	393	23.9	408	17	AI168396	AI168396 qa25d04.s
26	391.8	23.8	407	23	AI634093	AI634093 wa87g06.x
27	391	23.5	385	16	AI093060	AI093060 qa96f08.x
28	385.8	23.5	332	8	AA478614	AA478614 zv19c10.s
29	384	23.4	338	16	AI095032	AI095032 qa18g04.x
30	363	22.1	484	141	HI4168	HI4168 ym62g01.r1
31	348.4	21.2	412	5	AA303602	AA303602 EST16312
32	325.2	19.8	434	146	W25618	W25618 zb71d08.r1
33	324.2	19.7	361	16	AI144123	AI144123 qb62b08.x
34	320	19.5	341	18	AI277547	AI277547 q197c06.x
35	300	18.3	375	142	NS4182	NS4182 yx77c03.r1
36	277.8	16.9	307	15	AI074646	AI074646 ox81g11.s
37	275.6	16.8	481	142	N20280	N20280 yx42f05.s1
38	274	16.7	323	23	AI675296	AI675296 wc21f06.x
39	256.8	15.6	263	140	F17157	F17157 HSPD01678.H
40	255.6	15.6	269	6	AA385051	AA385051 EST96645
41	233	14.2	235	19	AI383176	AI383176 tc44f05.x
42	226.2	13.8	239	141	F35345	F35345 HSPD34146.H
43	225.2	13.7	254	141	F32964	F32964 HSPD34092.H
44	165.2	10.1	183	143	NS7623	NS7623 y218g05.r1
45	145	8.8	535	19	AI357592	AI357592 gy14h05.x

## ALIGNMENTS

RESULT 1  
 BE856748/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE

BE856748 586 bp mRNA  
 7f68b12.x1 Soares NSF\_F8\_9M\_OT\_PA\_P\_S1 Homo sapiens cDNA clone  
 IMAGE:3299807 3', mRNA sequence.  
 BE856748  
 BE856748.1 GI:10370067  
 EST.  
 human.

29-SEP-2000

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 586)  
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.jnl.gov) for further information.  
Seq primer: -400p from Gdbco  
High quality sequence stop: 444.  
FEATURES  
Source location/Qualifiers  
1..586  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:329807"  
/clone\_lib="Soares\_NSF\_F8\_9W\_OT\_PA\_P\_S1"  
/lab\_host="DH10B"  
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with  
a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;  
Equal amounts of plasmid DNA from five normalized  
libraries were mixed, and ss circles were made in vitro.  
Following HAP purification, this DNA was used as tracer in  
a subtractive hybridization reaction. The driver was  
PCR-amplified cDNAs from pools of 5,000 clones made from  
the same 5 libraries. The pools consisted of the following  
libraries and clones: Soares NBHF pool 1:  
309384-310919, 323208-325895 Soares NB2HP pool 1:  
145032-147335, 147720-148103, 148872-149255, 15002 -  
150407, 151176-152327 Soares NB2HB-9W pool 1:  
758280-760583, 772104-774407 Soares NBHPA pool 1:  
304776-306311, 320136-322823, 326280-326663 Soares NBHOT  
pool 1: 723720-726407, 739080-740999 Subtraction by Bento  
Soares and M. Fatima Bonaldo."

BASE COUNT 163 a 140 c 121 g 162 t  
ORIGIN

Query Match 35.3%; Score 579.6; DB 136; Length 586;  
Best Local Similarity 99.3%; Pred. No. 2e-132;  
Matches 582; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1056 gataaatttgtagctgtgacactctgctgtgtctgtaattcagctgtccagctg 1115  
|||||  
DB 586 GATTAATTTTGTAGCTGTGACACTCTGCGCTGGCTCTGTAATTCAGCTGCACGATG 527

QY 1116 gctgactgataatgagacagctctcatctgacccaactcttcctccactgaagctctc 1175  
|||||  
DB 526 GCTGACTGATGAAATGGACACGCTCATCTGACCCACTCTTCCTCCACTGAAGGCTCTTC 467

QY 1176 acgggacctccagtgagacaaaggatgacagcgcgctcgcatgccccaggaagcagctca 1235  
|||||  
DB 466 ACGGCGCTCCAGCTGACCAAAAGGATGCAAGCGCGCTCGATGCGCCAGGCGCAGCTCA 407

QY 1236 agagttccaaagatctcagatttggttttagtcatgatatataaacaagctctcaactcg 1295  
|||||  
DB 406 AAGAGTTCCAAAGATCTCAGATTGTTTGTAGCATGAATACATAAACAATCTCAAACTCG 347

QY 1296 ccaaatctttccctcccttttgaagcgaactgggccaatttggtgtaagagtggtgag 1355  
|||||  
DB 346 CACAATTTTTCCTCCCTTTGAAAGCCACTGGGCGCAATTTGTGTTAAGAGTGTGTAG 287

QY 1356 atcaagaatgtagacatcttgccagtggttgcagaagaatccaagagtgattggtct 1415  
|||||  
DB 286 ATTAAGAAGTGAAGTGAAGTCTTTGCGCAGTGTCTGAGAAAGATCCAGAGGATTTGGCT 227

QY 1416 tagttgtaaggagcttaagatcagagctgtaataatgagacaagtgggccagcttagcatc 1475  
|||||  
DB 226 TAGTTGTAAAGGCTTTAGAGTACAGGCGCAATATGAGGACAAGTGGGCGCAGCTTAGCATC 167

QY 1476 tgcagagatcaatctgagagctctgttctgcatctccaccagagagctagctctga 1535  
|||||  
DB 166 TGCAGAGATCAATCTGAGAGCTTCTGTCATCTGCGACGAGAGCTAGGCTCTTGA 107

QY 1536 tcttctcttagattgaagctgtctctgacacaaattattgttaaagttagtagtc 1595  
|||||  
DB 106 TCTTCTCTTAGATTGAAGTGTGCTGTCGACACAAATATTGTGAAGTTGAAGTTGC 47

QY 1596 ttttttaaatcattaaagaagctgtcgtgaaaaaataaaaaa 1641  
|||||  
DB 46 TTTTAAATCAATTAAAGAGGCTGCTGACGAAAAA 1

RESULT 2  
AM956791 579 bp mRNA EST 01-JUN-2000  
LOCUS AM956791 MAGE resequences, MAGD Homo sapiens cDNA, mRNA sequence.  
DEFINITION AM956791  
ACCESSION AM956791.1 GI:8146474  
VERSION  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 579)  
AUTHORS Hegde,P., Qi,R., Abernathy,K., Dharp,S., Gaspard,R., Gay,C., Holt  
I.E., Saeed,A.I., Sharov,V., Lee,N.H., Yeatman,T.J. and  
Quackenbush,J.  
TITLE Assessment of gene expression patterns in a model of colon tumor  
metastasis using a 19,200 element cDNA microarray  
JOURNAL Unpublished (2000)  
COMMENT Contact: John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: johnq@igr.org  
Plate: 96  
Seq primer: Reverse.  
FEATURES  
Source location/Qualifiers  
1..579  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_lib="MAGE resequences, MAGD"  
/note="Vector: pBluescriptSkm"  
BASE COUNT 168 a 117 c 121 g 171 t 2 others  
ORIGIN

Query Match 33.2%; Score 545.4; DB 97; Length 579;  
Best Local Similarity 98.4%; Pred. No. 5.5e-124;  
Matches 560; Conservative 0; Mismatches 8; Indels 1; Gaps 1;

QY 872 aagcaatctaaataataataagatagcgactatatactactctttaaatacaactgt 931  
|||||  
DB 1 AAGCAATCTAAATAATAATAATATATAGTATAGCACTATATACCTTAATAATCAACTGT 60

QY 932 ttgataaagcagagcatttlatatatacaatgaagactctgttcaattctt 991  
|||||  
DB 61 TTTGATATAGCAGCAGCATTTTATATATCAAAAGAGACTACTCTTTACATTTCTT 120

QY 992 aacataaacaatcglttttaactctctctgtgtagaatttttaagaacaaatlggaac 1051  
|||||  
DB 121 AACATATAAACATCGTTTCTTCTCTGCTAGCAATTTTAAAGCATTAATGGAATC 180

QY 1052 ctgtgataaatttttagctgtgtaactctgctgtgggtctgtgaattcagcctgacc 1111  
|||||  
DB 181 CTTGATTAATTTTGTAGCTGTGACACTCTGCGCTGGGTCTGTGAATTCAGCTGTGCAC 240

QY 1112 gatgctgactatgaatgagacagctctcatctgacccaactctccctcaactgaagt 1171  
|||||  
DB 241 GATGCTGACTATGAATGAAGACAGCTCATCTGACCCACTCTTCTTCCACTAAGGT 300

QY 1172 ctccacggcctccaggctggacccaaggatgcacacggcgctcgcacatgccccaggcgca 1231  
|||||  
Db 301 CTTCACGGGCTCCAGGTGACAAAGGATGCACAGGGGCTCGCATGCCCCAGGGCGCA 360  
QY 1232 gctcaaggtcccaagatctcagatcttggttttgatcatgataataaacagctctaaa 1231  
|||||  
Db 361 GCTAAGAGTTCACAAAGATCTCAGATTGGTTTGTATGATCATGATTAACATGATCTCAAA 420  
QY 1292 ctccacaaattttccccccttgaagacacacgctggggccatttgtgttaagagtg 1351  
|||||  
Db 421 CTGCACAAATTTTCCCTCCCTTTGAAAAGCAGCTGGGCGCAATTTGGTTT-AAAGGTGG 479  
QY 1352 tggataagaagtggaacgttacatcttggcagttgctcagaagaatcccaagcagatatt 1411  
|||||  
Db 480 TGAGATTAAGAGTGAAGCTGACATCTTTGGCCAGTGCAGAAAGATCAACAGATATT 539  
QY 1412 ggccttaagttaagggcttaagatcagg 1440  
|||||  
Db 540 GGCCTAATTGTAAAGGCTTATGATCAAG 568

RESULT 3  
AI870335/c 529 bp mRNA EST 07-MAR-2000  
LOCUS w172112.x1 NCI\_CGAP\_Brn25 Homo sapiens cDNA clone IMAGE:2430479 3',  
DEFINITION mRNA sequence.  
ACCESSION AI870335  
VERSION AI870335.1 GI:5544303  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 529)  
AUTHORS NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute / National Institute of Neurological  
Disorders and Stroke, Brain Tumor Genome Anatomy Project  
(CGAP/BRGAP), Tumor Gene Index  
JOURNAL Unpublished (1998)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,  
Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima  
Bonaldo, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
found through the I.M.A.G.E. Consortium/LINL at:  
www-bio.liln.gov/bbrp/image/image.html  
Insert length: 567 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 455.  
Location/Qualifiers

## FEATURES

source

1. 529  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone\_id="NCI\_CGAP\_Brn25"  
/tissue\_type="anaplastic oligodendroglioma"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pTZ19-Pac (Pharmacia) with a  
modified polylinker. Site\_1: Not I; Site\_2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer [5'  
TGTTCACATCTGAAGTGAAGGAGGCGCGATGAGTTTGTGTGTGTGTGTGT  
T 3']; double-stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified pTZ19 vector.  
Library is normalized, and was constructed by Bento  
Soares and M. Fatima Bonaldo."  
BASE COUNT 149 a 129 c 112 g 139 t

## ORIGIN

Query Match 32.1%; Score 527; DB 26; Length 529;  
Best Local Similarity 100.0%; Pred. No. 1.9e-119;  
Matches 527; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1100 caagctgctccacgagctgctgactgatatgaatggacacgctctcactcct 1159  
|||||  
Db 529 CAGGCTGTACCGATGGCTGACTGATGAATMGACAGCTGTCAATGACCCACTTCT 470  
QY 1160 tccactgaaggtcttcacgggctccaggtgagacaaaggatgacaggcgctcgcat 1219  
|||||  
Db 469 TCACACTAAAGCTTTCACAGGGCCCTCCAGTGGACAAAGGATGACAGCGGCTCCAT 410  
QY 1220 gccccaggcgacagctaaagattcccaagatcctcagatttggttttaactgatacata 1279  
|||||  
Db 409 GCCCAGGGCCAGCTTAAGAGTTCCTCAAGATCTCAGATTGTGTTTACTCATGATACATA 350  
QY 1280 aacagctcacaactcgacaaattttccccccttgaagacacactgaggccaattgtg 1339  
|||||  
Db 349 AACAGTCTCAAACTCGACAAATTTTCCCTTTTGAAGCCACCTGGGCGCAATTTGTG 290  
QY 1340 gtaagaagtggtgagataaagaagtggaacgtgacatccttgcagttgccaagaatc 1399  
|||||  
Db 289 GTTAAGAGGTGCTGATGATGAAGATGGAAGTGAACATCTTGGCAGTGTGCAGAAAGATC 230  
QY 1400 caagcaagttattgcttaagtttgtaagggctttagatcagcgctgaatatgaagcaaat 1459  
|||||  
Db 229 CAACGAGTATGTGGCTTAAGTTGTAAAGGCTTTAGATCAGCTGAATATAGCAAAAT 170  
QY 1460 gggcagcttagacatctgcagagatcaatctgagggcttctgttcacatctgcacag 1519  
|||||  
Db 169 GGGCAGCTTAGACATCTGCGAGATCAATCTGAGGCTTCTGTTTCGATCTGCGCAGC 110  
QY 1520 agagctaggtccttgatccttctttagatgaaagtcgtctctgacaacattattg 1579  
|||||  
Db 109 AGACCTAGGCTCTGATCTTTCTTTGATGTAAGTAAGTCTGCTGAAACAAATATTG 50  
QY 1580 taaagttagttagttctttttaaatacatcaaaagagctgctgaa 1626  
|||||  
Db 49 TAAAGTTAGTATGTTCTTTTAAATCATTAAGAGGCTTGCTGAA 3

RESULT 4  
BE465289/c 522 bp mRNA EST 27-JUL-2000  
LOCUS hw14c08.x1 NCI\_CGAP\_Lu24 Homo sapiens cDNA clone IMAGE:3182894 3',  
DEFINITION mRNA sequence.  
ACCESSION BE465289  
VERSION BE465289.1 GI:9511064  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 522)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
found through the I.M.A.G.E. Consortium/LINL, send email to:  
info@image.liln.gov  
Seq primer: -40UP from Gibco



High quality sequence stop: 470.

FEATURES  
Location/Qualifiers  
1. 522

Source  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3182894"  
/clone.lib="NCI\_CGAP\_Lu24"  
/tissue.type="carcinoid"  
/lab.host="DH10B"  
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker. Plasmid DNA from the normalized library NCI\_CGAP\_Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids 141420-141791 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 147 a 125 c 106 g 144 t  
ORIGIN

Query Match 31.5%; Score 517.2; DB 108; Length 522;  
Best Local Similarity 99.4%; Pred. No. 4.9e-117;  
Matches 519; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1116 gcttgatgtgtaaatggaacgtctcatctgacccactcttcctccactgaaggtcttc 1175  
|||  
Db 522 GCGATGATGAAATGGACACGCTCATGTGACCCACTCTTCCTCCACTGAAGGCTTTC 463  
|||  
Oy 1176 acgggctccaggtgacccaagaagatgcacagggcgctgcagcccgagggcagacta 1235  
|||  
Db 462 ACGGGCTCCAGGTGACCAAGAGGATGCACAGCGGCTCGCTGCCAGGCGCAGCTA 403  
|||  
Oy 1236 agagttccaaagatctcagattggtttagtcataatcaataaagttcaactcg 1295  
|||  
Db 402 AGAGTCCAAAGATCTCAGATTGTTTATGATCATGATATCAATAAAGTCTCAAACTCG 343  
|||  
Oy 1296 caaatcttccccccttgaagccactggggccaattgtgtttaagaggtgtgtgag 1355  
|||  
Db 342 CCAATTTTTCCTCCCTTTGAAAGCCACTGGGCGCAATTTGGTAAAGAGTGTGTGAG 283  
|||  
Oy 1356 ataagaagtggaagctgacatctcttcaggttgtagaagaatccaaagcaggtatgtgct 1415  
|||  
Db 282 ATTAAGAAGTGAAGCTGATCTTTCCAGTGTCTGCAAGATCCAAAGAGTATTTGGCT 223  
|||  
Oy 1416 tagttgtaaggccttaggaatcagctgaataatgaagacaagtgggccagcttagcacc 1475  
|||  
Db 222 TAGTGTGAAGGCTTTAGATCAGGCTGAATATGAGGACAAAGTGGCCACGTTTACATTC 163  
|||  
Oy 1476 tgcagagatcaatctggaggtctctgtctgcattctgcacagagagcttagtctctta 1535  
|||  
Db 162 TCGAGAGATCAATCTGAGAGCTTCTTCTGCTGCTCCAGAGAGCTAGGCTCTTGA 103  
|||  
Oy 1536 tctttctttagatgaagctgtctctgcgaacaattattgttaaaagttagttagtc 1595  
|||  
Db 102 TCTTTCTTTAATTAAGTCTGCTCTGCAACACAAATTTATTGTAAGTATGTTAGTTC 43  
|||  
Oy 1596 tttttaaatcaataaagaagctgtctgaataaaaaa 1637  
|||  
Db 42 TTTTAAATCAATTAAGAGGCTTCTGAGAGATTAATAAAA 1  
|||

RESULT 5  
AW015888 503 bp mRNA EST 10-SEP-1999  
LOCUS AW015888/C  
DEFINITION UI-H-B10-aa0-d-11-0-UI-s1 NCI\_CGAP\_Sub1 Homo sapiens cDNA clone  
IMAGE:2709956 3', mRNA sequence.  
ACCESSION AW015888  
VERSION AW015888.1 GI:5864645  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS  
TITLE

JOURNAL  
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 503)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.Strausberg@nih.gov  
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA library Preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: [www.bio.livl.gov/db/rp/image/image.html](http://www.bio.livl.gov/db/rp/image/image.html)  
Seq primer: M13 Forward  
POLYA=Yes.

FEATURES  
Source  
Location/Qualifiers  
1. 503  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2709956"  
/clone.lib="NCI\_CGAP\_Sub1"  
/lab.host="DH10B (Life Technologies)"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site\_1: Not I; Site\_2: Eco RI; The NCI\_CGAP\_Sub1 library is a subtracted library derived from BI. BI constitutes a mixture of 21 normalized or subtracted NCI\_CGAP libraries: NCI\_CGAP\_Co4, NCI\_CGAP\_Pr22, NCI\_CGAP\_Pr28, NCI\_CGAP\_Co10, NCI\_CGAP\_Co16, NCI\_CGAP\_Kid5, NCI\_CGAP\_Kid12, NCI\_CGAP\_Kid3, NCI\_CGAP\_Kid11, NCI\_CGAP\_Lym2, NCI\_CGAP\_Br2, NCI\_CGAP\_Co8, NCI\_CGAP\_Co11, NCI\_CGAP\_Lym2, NCI\_CGAP\_Br23, NCI\_CGAP\_Lu5, NCI\_CGAP\_Lu24, NCI\_CGAP\_Lu19, NCI\_CGAP\_GC4, NCI\_CGAP\_GC6, NCI\_CGAP\_Br25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below: NCI\_CGAP\_Kid3 pool 1 L1AM 3334-3337, 3682-3683, 3798-3803 (IMAGE Clonoids 1322376-1323911, 1456008-1456751, 1500552-1502855) NCI\_CGAP\_Kid5 pool 1 L1AM 3338-3342, 3722-3725, 3776-3778 (IMAGE Clonoids 1323912-1325831, 1471368-1472903, 1492104-1493253) NCI\_CGAP\_Lu5 pool 1 L1AM 3575-3582, 3851-3854 (IMAGE Clonoids 1414920-1417991, 1520904-1522439) NCI\_CGAP\_GC4 pool 1 L1AM 3164-3167, 3716-3720, 3733-3735 (IMAGE Clonoids 1257096-1258631, 1469064-1470983, 1475592-1476743) NCI\_CGAP\_Pr22 pool 1 L1AM 2457-2459, 2758-2759, 3062-3068 (IMAGE Clonoids 985608-986759, 1101192-1101959, 1217928-1220615) NCI\_CGAP\_Co10 pool 1 L1AM 2644-2653, 2871-2872 (IMAGE Clonoids 1057416-1061255, 1145084-1145351) The resulting subtracted library contained 530,000 recombinants. Subtraction was performed as previously described (Bonaldo Lennon & Soares (1996): Normalization and Subtraction: Two Approaches to Facilitate Gene Discovery. Genome Research 6, 791-806.  
TAG\_LIB=NCI\_CGAP\_Lu5  
TAG\_TISSUE=Lung  
TAG\_SEQ=CAAC"

BASE COUNT 139 a 118 c 101 g 145 t  
ORIGIN

Query Match 29.3%; Score 480.6; DB 38; Length 503;  
Best Local Similarity 98.2%; Pred. No. 5.2e-108;  
Matches 486; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Oy 1148 cccactctctccactgaaggtcttcacggcgctccaggtgacccaaggtgacaca 1207  
|||  
Db 498 CCACCTCTCTCCAGTGAAGGTCTTCCAGGCGCTCCAGAGGATGATCA 439  
|||

```

QY 1208 gggcgctcgatgcccccagggcagctaaagctccaaagctcagcttggctttagt 1267
| | | | |
Db 438 GGGCGCTCGCATGGCCCGAGGCGCAGTAAAGATTCCAAAGATCTCAGATTGGTTAGT 379
QY 1268 catgaatacaataaagctctaaactcgcaatttttcccccttttgaagccactgg 1327
| | | | |
Db 378 CATGAATACATAAACAGCTCAAACTCGCACAAATTTTCCCTTTTGAAGCCACTGG 319
QY 1328 ggcgaattgtgttaagaggtgtgagataaagatggaacgtgacatcttgcagtt 1387
| | | | |
Db 318 GGCCAAATTTGGTTAAGAGGTGTGAGATTAAGAAGTGAAGTGAAGTCTTTCAGATT 259
QY 1388 gtccagaagaatccaaagcaggtatgtgcttaagaggtcttaagatcagctgaata 1447
| | | | |
Db 258 GTCCAGAAATCCAAAGCAGGTATGTGGCTTAAAGGCTTAAAGGCTTAAAGTGAATA 199
QY 1448 tgaagaaagtgagcagcttaagcagctgagagatcagatcagcttggagctcttctg 1507
| | | | |
Db 198 TGAGGACAAATGGGCGCCAGTTAGCATCTGAGAGATCAATCTGGAGGCTTCTGTCTG 139
QY 1508 catctgcacagagagcttagtcttctgacatcttcttaagattgaagctctctcga 1567
| | | | |
Db 138 CATCTGCACAGAGAGCTGCTTGTATCTTTCTTTAGACTGAAGTCTCTCTGAA 79
QY 1568 cacaattattgtaaagctagctcttctttaaatacaataaagagctgtcga 1627
| | | | |
Db 78 CACAATATTATTGTAAAGTATAGTCTTTTAAATCATTAAGAGGCTGTCTGATG 19
QY 1628 aaaaaaaaaaaaaa 1642
| | | | |
Db 18 GAAAAAAAAAAAAA 4

RESULT 6
LOCUS A1817213 476 bp mRNA EST 07-MAR-2000
DEFINITION w62h12.x1 Soares_NSF_F8_9W_OT_PA_P.S1 Homo sapiens cDNA clone
IMAGE:2369735 3, mRNA sequence.
ACCESSION A1817213
VERSION A1817213.1 GI:5436281
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 476)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (InfoImage.lnl.gov) for further information.
Insert Length: 574 Std Error: 0.00
Seq primer: -40UP from Gldco
High quality sequence stop: 461.
Location/Qualifiers
1..476
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2369735"
/clone_id="Soares_NSF_F8_9W_OT_PA_P.S1"
/lab_host="DH10B"
/note="Organ: pooled; Vector: p773D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following

```

```

BASE COUNT 132 a 116 c 95 g 132 t 1 others
ORIGIN
libraries and cloneIDs: Soares NBHSF pool 1:
309384-310919, 323208-325895 Soares NB2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares NB2HF8-9W pool 1:
758280-760583, 772104-774407 Soares NBHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NBHOT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."

Query Match 28.4%; Score 467; DB 25; Length 476;
Best Local Similarity 98.7%; Pred. No. 1.2e-104;
Matches 470; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1158 ctccactgaaggtcttcacgggctccagcttgacccaaggatgacaggcgctcgc 1217
| | | | |
Db 476 CTTCACATGAAGGCTTCACGGGCTCCAGGTGACCAAGGATGACAGCGCTCGC 417
QY 1218 atgcccaaggccagctaaagatccaaagatccagattgtgttagcatgaataca 1277
| | | | |
Db 416 ATGCCCAAGGCCAGCTAAAGATTCAAAAGATCTCAGATTGGTTAGTCATGAATACA 357
QY 1278 taacagctcaaacctgcacaattttccccctttgaagccactgggccaatttg 1337
| | | | |
Db 356 TAAACAGTCTCAAACTGCACAAATTTTCCCTTTTGAAGCCACTGGGGCCAAATTG 297
QY 1338 tgtgttaagaggtgtgtgataaagaatggaacgtgacatcttggccagctgtcgaagaa 1397
| | | | |
Db 296 TGTGTTAAGAGGTGTGAGATTAANAAGTGAACGTGAACATCTTGGCCAGTTGTGGAAGAA 237
QY 1398 tccaagcaggtatgtgcttgaaggtcttaaggtatcagctgatatgaagacaa 1457
| | | | |
Db 236 TCCAAGCAGAGATTGGCTTACTTACTTAAGGCTTTAGATCAGCGCCATATGAGAGAAAA 177
QY 1458 gtgggccaagcttagcatctgcagaatcaatctggaagctctgttctgcattcgca 1517
| | | | |
Db 176 GTGGGCCACGTTAGCATCTGCAGAGATCAATCTGGAGGCTTGTCTGCATCTGCCA 117
QY 1518 cgaagctaggtcttcttcttcttcttgaatgaagctgtctctgaaacaattatc 1577
| | | | |
Db 116 CGAGAGCTAGGCTCTTGATCTTTCTTTAGATTAAGATGCTGCTCGAACACATATTAT 57
QY 1578 tgtaaagttagtgctcttctttaaatacaataaagagctgtctaaaaaaaa 1633
| | | | |
Db 56 TGTAAAGTTAGAGCTCTTTTAAATCATTAAGAAGCTTCTGACAGAAAAA 1

RESULT 7
LOCUS AA778133 461 bp mRNA EST 05-FEB-1998
DEFINITION zf46h06.s1 Soares_fetal_heart_NBH19W Homo sapiens cDNA clone
IMAGE:380027 3, mRNA sequence.
ACCESSION AA778133
VERSION AA778133.1 GI:2837534
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 461)
Hillier L., Allen M., Bowles L., Dubuque T., Geisels G., Jost S.,
Kizman D., Kucaba T., Lacy M., Le N., Lennon G., Mair M., Martin
J., Moore B., Schellenberg K., Steptoe K., Tan F., Theising B.,
White Y., Wylie T., Waterston R. and Wilson R.
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

```



QY 1475 ctgcagagatacaatctgaggctctgtctgtcgtcattctgcacagagagctagtccttg 1534  
|||||  
Db 187 CTGAGAGATCAATTTGAGAGGCTTCTGTTTGCATTCCTGCCCGAGAGCTAGAGTCTTG 128  
QY 1535 atctctctcttagatgtgaagtctgtctgtcgaacacattattgttaaaagttagt 1594  
|||||  
Db 127 ATCTTTCTTTAGATGGAAGCTCTCTCTGTAACCCCAATTAATTTGTAAGTTAGTAGT 68  
QY 1595 ctttttaacatcaaaagagctgtctgtgaataaaaaa 1642  
|||||  
Db 67 CTTTTTAAATCATTAATAAGAGGCTTGCTGAAGATTAATAAAAAA 20

RESULT 9  
LOCUS N52801 466 bp mRNA EST 15-FEB-1996  
DEFINITION y218g05.s1 Soares multiple\_sclerosis\_2NBHMSR Homo sapiens CDNA  
clone IMAGE:283448 3', mRNA sequence.  
ACCESSION N52801  
VERSION N52801.1 GI:1193967  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 466)  
AUTHORS Hallier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman  
,M., Hulman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,  
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaaskis,E., Waterston  
,R., Williamson,A., Woldmann,P. and Wilson,R.

TITLE The Washu-Merck EST Project  
JOURNAL Unpublished (1995)  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: est@wustl.wustl.edu  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: ml3 -40 forward  
High quality sequence stop: 467.

FEATURES  
Location/Qualifiers  
1..466

/organism="Homo sapiens"  
/db\_xref="GDB:3901976"  
/db\_xref="taxon:9606"  
/clone="IMAGE:283448"  
/clone\_lib="Soares\_multiple\_sclerosis\_2NBHMSR"  
/sex="male"  
/tissue\_type="multiple sclerosis lesions"  
/dev\_stage="Age 46"  
/lab\_host="DH10B (ampicillin resistant)"  
/note="Vector: pT73D (Pharmacia) with a modified  
polylinker V-type; phagemid; Site\_1: Not I; Site\_2: Eco RI  
; 1st strand CDNA was primed with a Not I - oligo(dT)  
primer [5',  
TGTACCAATCTGAGTGGAGCGCCGCAATTTTTTTTTTTTTTTT 3'],  
double-stranded CDNA was size selected, ligated to Eco RI  
adapters (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified pT73 vector  
(Pharmacia). Library went through one round of  
normalization to a Cot = 5. Library constructed by Bento  
Soares and M. Fatima Bonaldo. RNA from 4 multiple sclerosis  
lesions from one patient was kindly provided by Dr. Kevin  
G. Becker (NINDS/NIH)."  
BASE COUNT 131 a 117 c 91 g 127 t  
ORIGIN

Query Match 27.3%; Score 448.8; DB 143; Length 466;  
Best Local Similarity 99.4%; Pred. No. 3.5e-100;  
Matches 461; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1164 ctgaaggtctcagcggcctcccaagtgtgaccaaaggatgcacaggcgtcgtcgtcc 1223  
|||||  
Db 466 CTGAAGGCTCTTCACGGGGCTCCAGGTGGAGCCAAAGGAGTGCACAGGGGCTCGCATGCC 407  
QY 1224 cagggcagcgaagaagttccaaagatcagaattgtgttagtcaatgaaataaaca 1283  
|||||  
Db 406 CAGGGCCAGCTAAGAGTTCACAAAGATCTCAGATTGGTTAGTCATGAAATCAATTAACA 347  
QY 1284 gtctcaaatctgcacaatttttcccttttgaaagccactggggccaatttggttta 1343  
|||||  
Db 346 GTCGAACTCTGCACAAATTTTTCCTTTGAAAGCCACTGGGGCCCAATTTGTGTTA 287  
QY 1344 aagagtggtgataaagaagtggaacgtgacatcttgcagttgtcagaagaatccaaag 1403  
|||||  
Db 286 AGAGGTGTGAGATPAGAAGTGAACGTGACATCTTTCAGATTGTGTCAGAAATCCAAAG 227  
QY 1404 caggtattggtctgtgttaagggtcttaagatcagaagctgaatatagagacaagtg 1463  
|||||  
Db 226 CAGGATTTGGCTTGTGTTAAGAGGCTTTAGGATCAGGCCGGAATATGAGACAAAGTGGCC 167  
QY 1464 cac-gttagcatctgcaagaatcaatctggaggctctgttctgcattctgcacagaga 1522  
|||||  
Db 166 CACAGTTAGCATCTGCAGAGATCAATCTGGAGGCTTCTGTTCTGCATTCTGCCACGAGA 107  
QY 1523 gctaggtcctgattcttctttagatgaaagtcgtctctgacaacaattattgtaa 1582  
|||||  
Db 106 GCTAGGCTCTTGATCTTCTTTTATGATGAAGTCTGCTCTGCAACCAATATTATTGTA 47  
QY 1583 aagttagtagtctttttaaatacaataaagaagctgtgtgaa 1626  
|||||  
Db 46 AAGTTAGAAGTCTTTTAAATCATTAATAAGAGCTTGCTGAA 3

RESULT 10  
LOCUS AM614686/c 453 bp mRNA EST 23-MAR-2000  
DEFINITION h31b08.x1 NCI\_CGAP\_Lu24 Homo sapiens CDNA clone IMAGE:2956695 3',  
mRNA sequence.  
ACCESSION AM614686  
VERSION AM614686.1 GI:7319872  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1 (bases 1 to 453)  
AUTHORS NCI\_CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index

JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550

Email: Robert\_Strausberg@nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
image.llnl.gov/image/hcml/iresources.shtml  
Seq primer: -40UP from GIBCO  
High quality sequence stop: 448.

FEATURES  
Location/Qualifiers  
1..453

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2956695"  
/clone\_lib="NCI\_CGAP\_Lu24"  
/tissue\_type="carcinoid"  
/lab\_host="DH10B"  
/note="Organ: Lung; Vector: pT73D-Pac (Pharmacia) with a

modified polylinker; plasmid DNA from the normalized library NCI CGAP.Lu5 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 129 a 111 c 83 g 130 t

ORIGIN

Query Match 27.3%; Score 448.2; DB 93; Length 453;  
Best Local Similarity 99.3%; Pred. No. 4.9e-100;  
Matches 450; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1185 caggtgaccacaaaggatgacacagcgctgcagatgccacaggccagctaaagttcca 1244  
DB 453 CAGGTGACCAAAAGGATGACACAGCGCTGCAGTCCAGGCGCCAGTAAAGAGTTCCA 394  
QY 1245 aagatcagattgtgttagtcatgaatacaataaagctccaactgcacaaattt 1304  
DB 393 AAGATCTCAAGATTGTGTGTAGTCATGAATACATAAACAGTCCAACTGCACAAATTTT 334  
QY 1305 ttccccccttgaagacactggggccaatttggtgttaagagtggtgtgaataagagt 1364  
DB 333 TTCCCCCTTTGAAGACACTGGGGCCAATTTGTGTTAAGAGTGGTGATTAAGAGT 274  
QY 1365 ggaacgtgacatcttgcagctgtcagaagaatccaagaglatgtccttagttgaa 1424  
DB 273 GGAACGTGACATCTTGCAGTGTGTGAGAAATCCAAAGAGGATATGGCTTGTGTTAA 214  
QY 1425 gggctttagatcagctgaataatgagacaaagtggcgacgttagatctcagagat 1484  
DB 213 GGGCTTTAGATCAGGCTGATATGAGACAAAGTGGGCACGTTACATCTGACAGAGAT 154  
QY 1485 caatctggaagctctgttcttcgcatctgcacgaagagtagtctcttagcttctt 1544  
DB 153 CAATCTGAGAGCTTCTGTCTGATTTCTGCCAGAGAGCTAGTCTTATCTTTTCTT 94  
QY 1545 tagatggaagctctgtctcgaacacaaatatttgaagaagtagtattctttttaa 1604  
DB 93 TAGATTGAAGCTGTCTGTGACACAAATATTGTAAAGTAGTAGTCTTTTAA 34  
QY 1605 tcattaagagagctgtctgtaaaaaa 1637  
DB 33 TCATTAAAGAGGCTTCTGTAAGATTAATAAAAA 1

RESULT 11  
A1569688 446 bp mRNA EST 14-APR-1999  
LOCUS t004e05.x1 NCI.CGAP\_Brn25 Homo sapiens cDNA clone IMAGE:216656 3',  
DEFINITION mRNA sequence.  
ACCESSION A1569688  
VERSION A1569688.1 GI:4533062  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 446)  
NCI/NIHNS-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute / National Institute of Neurological  
Disorders and Stroke, Brain Tumor Genome Anatomy Project  
(CGAP/BRGAP), Tumor Gene Index  
Unpublished (1998)  
Contact: Robert Strausberg, Ph.D.  
Tel: (301) 496-1550  
Email: Robert.L.Strausberg@nih.gov  
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,  
Ph.D.  
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima

Bonaldo, Ph.D.  
CDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone Distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
[www-bio.11nl.gov/bdrp/image/image.html](http://www-bio.11nl.gov/bdrp/image/image.html)  
Insert Length: 565 Std Error: 0.00  
Seq primer: -400P from Glibco  
High quality sequence stop: 444.  
Location/Qualifiers  
1. 446  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:216656"  
/clone\_11b="NCI.CGAP\_Brn25"  
/tissue\_type="anaplastic oligodendroglioma"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: p7773D-Pac (Pharmacia) with a  
modified polylinker; Site: 1: Not I; Site 2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer (5'  
TGTACCAATCTGAAGTGGAGGCGGCGCATAGTGTGTGTGTGTGTGTGTGTGTGTGT  
T 3'); double-stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified p7773 vector.  
Library is normalized, and was constructed by Bento  
Soares and M. Fatima Bonaldo."

BASE COUNT 129 a 112 c 84 g 121 t

ORIGIN

Query Match 26.9%; Score 442.4; DB 22; Length 446;  
Best Local Similarity 99.8%; Pred. No. 1.3e-98;  
Matches 443; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1183 tccagtggaacaaaggatgacacagcgctgcagatgccacaggccagctaaagattc 1242  
DB 446 TCCAGTGGACCAAGGAGATGACACAGCGGCTGCATGCCAGGCGCACATTAAGATTTC 387  
QY 1243 caaagatcagattgtgtttagtcaatacaataaagctccaactgcacaaatt 1302  
DB 386 CAAAGATCTCAGATTGTGTTAGTCATGAATACATMAACAGTCTCAAACTCCACAAAT 327  
QY 1303 tttccccccttgaagacacactggggccaatttggtgttaagagtggtgaaataagaa 1362  
DB 326 TTTTCCCTTTTGAAGCCACTGGGCGCAATTTGTGTTAAGAGTGGTAATTAAGAA 267  
QY 1363 gtggaacgtgacatcttgcagctgtcagaagaatccaagcagatattgtccttagt 1422  
DB 266 GTGGAAGCTGACATCTTGTGCAAGTGTGCAGAAAGATCAAGCAGATTTGGCTTAGTGT 207  
QY 1423 aagggtttagatcaggtcgaatataagagaacaagtggcgacgttagatctgcagag 1482  
DB 206 AAGGGCTTTAGATCAGGCTGATATGAGACAAAGTGGGCCACGTATGATCTGCAGAG 147  
QY 1483 atcaatctggaagctctgttcttcgcatctgcacgaagagtagtctcttagcttctt 1542  
DB 146 ATCAATCTGAGAGCTTCTGTCTGATTTCTGCACAGAGCTAGGCTTGATCTTTTTC 87  
QY 1543 tttagattgaagctgtctcgaacacaaatttgaagaagtagtattctttttaa 1602  
DB 86 TTTAGATTGAAGTCTGTCTGTGACACAAATTTTGTAAAGTTAGTACTTCTTTTAA 27  
QY 1603 aatcatnaaagagagctgtctgaa 1626  
DB 26 AATCATTTAAAGAGGCTGTCTGAA 3

RESULT 12  
A1199159 443 bp mRNA EST 01-DEC-1998  
LOCUS A1199159/c  
DEFINITION q146h07.x1 NCI.CGAP\_Brn25 Homo sapiens cDNA clone IMAGE:1859581 3',  
DEFINITION mRNA sequence.  
ACCESSION A1199159



OY	1163	actaaagctcttaacaggcgtccaaagtga - ccaaggagatcaaggcgcgtc - -cat	1213
Db	475	ACTAAGNCTTCTTCAGGCCCTTCAGNTGACCCAAAGGATCAACAGGCGGCTGCCAT	416
OY	1220	gccccaggccagctaaagatcccaaaagct - cagatttgcttttaagtcagaaatcat	1278
Db	415	GCCCCAGGCGCAGCTAAGATGTCCAAAGATCTCAGATTGTGGTTTATGATGAATACAT	356
OY	1279	aaacagctctaaactcgcaaatlttttcccccctttgaagccacctggggccaattgt	1338
Db	355	AAACAGCTCTMAACCTCCACAAATTTTTTCCCTTTTGAAACCACTGGGGCAATTTGT	296
OY	1339	ggttaagagtggtgagataagaagctggaacgtgacatcttgcacatgttcaagaagat	1398
Db	295	GCTTAAGAGTGGTGGAGATTAAGAAGTGGACAGTGACATCTTGGCCAGTTGTCAAGAAGAT	236
OY	1399	ccaagcagctatttgcttagctgtaaggccttaagatcagctgcaatataaggaacaag	1458
Db	235	CCAAGCAGGATTTGGCTTAAGTCTTAAGGCGCTTTAGATCAGCGCTGAATATGAGGCAAGAAG	176
OY	1459	tgggcacgcttaagatctgaaagatcaactctggagagctctggttctgcgaattcggcac	1518
Db	175	TGGGCACAGTTAGCATCTGCAGAGATCAATCTGGAGCGCTTGTTCTGCACATCTGCCAC	116
OY	1519	gagagctaggctcccttgatcttctcttaagatgaaagctgctctctgacaacaattatct	1578
Db	115	GAGAGCTAGGACCTTGATCTTTTCTTTAGATTGAAGAGTGTGCTCTGAAACCAATATATT	56
OY	1579	gtaaagcttgatgctctctttttaaactcaataaagaagctgctggaaaaaaaa	1633
Db	55	GTAATAAGTTAGTAGTCTCTTTTAAATCAATTAACAGGCTTGCTGGAAGGATATAA	1

RESULT 14  
 LOCUS AI371173/c  
 DEFINITION AI371173 425 bp mRNA  
 ID: c11.x1 Soares\_total\_tetus\_ND2HF8\_9w Homo sapiens cDNA clone  
 IMAGE: 2043572 3', mRNA sequence.  
 ACCESSION AI371173  
 VERSION AI371173.1 GI:414926  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 425)  
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP).  
 Tumor\_Gene\_Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Tel: (301) 496-1550  
 Email: Robert\_Strausberg@nih.gov  
 This clone is available royalty-free through LNL ; contact the  
 IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
 Insert length: 558 Std Error: 0.00  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 423.  
 Location/Qualifiers  
 1..425

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone IMAGE:2043572"
/clone_11b= Soares_total_fetus_Nb2HF8_9w"
/dev_stage= 8-9 weeks"
/lab_host="DH10B"
/notes="vector: pTY73D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand
was prepared from mRNA obtained from pooled 8-9 week
(totals) fetus material with a Not I - oligo(dT) primer [5'
TGTTTCACATTCGAAGTGGACGGCCGCTATATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors

```

(Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo. "

Query Match	25.68;	Score 420.2;	DB 19;	Length 425;
Best Local Similarity	99.38;	Pred. No. 3.9e-93;		
Matches 422;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;

QY	1211	ggctcgatbccccagggccagctaaagttccaaagatctagaattgatttagcat	1270
Db	425	GGCTCGATGCCCCAGGCGCAGCTAAGAGTCTCCAAAGATCTAGATTGGTTTAGTAT	366
QY	1271	gaatacataaacagtcctcaactcgcacaattttccccccttgaagccaatlyggc	1330
Db	365	GAATACATAAACAGTCTCAAACTCGCACAAATTTTTCCCCCTTTGAAGCCACATGGGCG	306
QY	1331	caatttggttaagaggtggtgagtaagaagtgaagaaogtaactccttgcagttgct	1390
Db	305	CAATTGTGTGTTAAGAGTGCTGATGAATGAAGTGAAGTGAACGTGACATCTTGGCAATTTCTC	246
QY	1391	agaagaaaccaagaogtattgcttgaattgaagaagccttgaatcagactgaatga	1450
Db	245	AGAAGATCCACAGAGGTATTTGGCTTAGTTAGTTGAAGGCTTTAGATCAGCGCTGATATGA	186
QY	1451	ggacaaagtyggccaogttagcaatctgcagagatacaactctgaagcctctgttctgat	1510
Db	185	GGACAAAGTGGCGACGCTGTAGCATCTGCAGAGATCAATCTGAAGGCTTCTGTTTCTGAT	126
QY	1511	tcctccacagagactaggtcttcctgatactttctttagatattgaagctcgtctctgaac	1570
Db	125	TCTCCACGAGAGCTAGGTCCTTGATCTTTCTTTAGATTGAAGTCTGCTCTGAAC	66
QY	1571	aattcattgtlaaagttagtagtctctttaaatacatlaaagaagccttgctgaaaaa	1630
Db	65	AATATTATTGTAAGAAGTTAGATTCTTTTAAATCATTAAGAAGGCTTGCTGAAGAT	6
QY	1631	aaaaa 1635	
Db	5	AAAAA 1	

RESULT	15
LOCUS	A1085287/c
DEFINITION	A1085287 419 bp mRNA EST 17-AUG-1998 ql18f01.x1 NCI_CGAP_Brn25 Homo sapiens cDNA clone IMAGE:1750393 3'
ACCESSION	A1085287
VERSION	A1085287
KEYWORDS	A1085287.1 GI:3423710
SOURCE	EST.
ORGANISM	human. Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 419) NCI/NINDS-CGAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> . National Cancer Institute / National Institute of Neurological Disorders and Stroke, Brain Tumor Genome Anatomy Project (CGAP/BTCAP), Tumor Gene Index Unpublished (1998)
JOURNAL	Contact: Robert Strausberg, Ph.D.
COMMENT	

Email: Robert\_Strausberg@nh.gov  
Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D., Ph.D.  
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:  
www.bio.livl.gov/bbrp/image/image.html  
Seq primer: -40m13 fwd. ET from Amersham  
High quality sequence stop: 392.

## FEATURES

## Source

1. .419  
Location/Qualifiers

/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1750393"  
/clone\_lib="NCI\_CGAP\_Brn25"  
/tissue\_type="anaplastic oligodendroglioma"  
/lab\_host="DH10B"  
/note="Organ: brain; Vector: pT73D-Pac (Pharmacia) with a  
modified po.ylinker; Site\_1: Not I; Site\_2: Eco RI; 1st  
strand cDNA was primed with a Not I - oligo(dT) primer [5'  
TCTTACCAATCCGAGAGTGGAGCGGCCCATAGCTTTTCTTTTCTTTTCTTTT  
T 3']; double-stranded cDNA was ligated to Eco RI  
adaptors (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of the modified pT73 vector.  
Library is normalized, and was constructed by Bento  
Soares and M. Fatima Bonaldo."

BASE COUNT 126 a 101 c 78 g 114 t  
ORIGIN

## Query Match

25.4%; Score 417; DB 16; Length 419;

Best Local Similarity 100.0%; Pred. No. 2.4e-92;  
Matches 417; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1210 cggctcgatgccccagggccagctagaagctccaaagatcctaagattggttagtca 1269
|||
DB 419 cggctcgatgccccagggccagctagaagctccaaagatcctaagattggttagtca 360
QY 1270 tgaatacataaacaagctcaactcgacaattttccccccttgaagcgaactgggg 1329
|||
DB 359 tgaatacataaacaagctcaactcgacaattttccccccttgaagcgaactgggg 300
QY 1330 ccaatttggttaagaagctggatgaataagaagctgaacgtgacactcttgccagtgt 1389
|||
DB 299 ccaatttggttaagaagctggatgaataagaagctgaacgtgacactcttgccagtgt 240
QY 1390 cagaagaatccaagcagctatcgcttagtctglaaggccttagatcagcagctgaatc 1449
|||
DB 239 cagaagaatccaagcagctatcgcttagtctglaaggccttagatcagcagctgaatc 180
QY 1450 aggaacaaagtgggcccgttaagcatctgacagaagatcaatcgaagctctgttctgca 1509
|||
DB 179 aggaacaaagtgggcccgttaagcatctgacagaagatcaatcgaagctctgttctgca 120
QY 1510 ttctgcacgagagctaggtcccttgactcttctttagattgaagctcgtctcgaaca 1569
|||
DB 119 ttctgcacgagagctaggtcccttgactcttctttagattgaagctcgtctcgaaca 60
QY 1570 caattattgtaaaagttagttagtcttctttaaalcataaagaagcgtcgtgaa 1626
|||
DB 59 caattattgtaaaagttagttagtcttctttaaalcataaagaagcgtcgtgaa 3
```

Search completed: February 24, 2001, 21:04:33  
Job time: 6986 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2001, 21:06:11 ; Search time 86.16 Seconds  
(Without alignments)  
3071.323 Million cell updates/sec

Title: US-09-320-713-3

Perfect score: 1642

Sequence: 1 ggaatcgcgcagcagctctg.....tgaaaaaaaaaaaaaaaaaa 1642

Scoring table: IDENTITY\_NDC  
Gapop 10.0 , Gapext 1.0

Searched: 280836 seqs, 80580151 residues

Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents, NA:\*

- 1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq:\*
- 3: /cgn2\_6/ptodata/2/ina/6.COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Query Length	DB ID	Description
1	56	3.4	3624	1	US-07-951-715A-6 Sequence 6, Appli
2	56	3.4	3624	2	US-08-459-448A-6 Sequence 6, Appli
3	56	3.4	3624	3	US-08-459-595A-6 Sequence 6, Appli
4	56	3.4	3624	3	US-08-459-504B-6 Sequence 6, Appli
5	56	3.4	3624	3	US-08-459-444-6 Sequence 6, Appli
6	56	3.4	3624	3	US-09-053-549-7 Sequence 7, Appli
7	53.4	3.3	2734	3	US-09-135-021-7 Sequence 7, Appli
8	53.4	3.3	3181	3	US-09-135-021-1 Sequence 1, Appli
9	53.4	3.3	3181	3	US-09-135-021-5 Sequence 5, Appli
10	52.6	3.2	1485	1	US-08-471-601-23 Sequence 23, Appli
11	52.6	3.2	1485	1	US-08-474-556-23 Sequence 23, Appli
12	52.6	3.2	1485	1	US-08-479-382-23 Sequence 23, Appli
13	52.6	3.2	1485	1	US-08-470-354-23 Sequence 23, Appli
14	52.6	3.2	1485	1	US-08-479-383-23 Sequence 23, Appli
15	52.6	3.2	1485	1	US-08-479-041-23 Sequence 23, Appli
16	52.6	3.2	1485	1	US-08-479-041-23 Sequence 23, Appli
17	51.4	3.1	1821	2	US-09-128-155-16 Sequence 16, Appli
18	51.4	3.1	1821	2	US-08-265-310-7 Sequence 7, Appli
19	51.4	3.1	1821	3	US-08-951-742-7 Sequence 7, Appli
20	50.8	3.1	2712	3	US-09-025-691-4 Sequence 4, Appli
21	50.6	3.1	8438	1	US-07-945-283-1 Sequence 1, Appli
22	50.4	3.1	1777	2	US-08-173-508-5 Sequence 5, Appli
23	50.4	3.1	1777	2	US-08-265-310-5 Sequence 5, Appli
24	50.4	3.1	1777	2	US-08-951-742-5 Sequence 5, Appli
25	50.4	3.1	2418	1	US-08-462-484-1 Sequence 1, Appli
26	50.4	3.1	2418	1	US-08-441-147-1 Sequence 1, Appli
27	49.6	3.0	624	4	PCT-US95-07536-1 Sequence 1, Appli
28	49.6	3.0	624	2	US-08-426-599B-3 Sequence 3, Appli
29	49.6	3.0	1820	1	US-08-173-508-7 Sequence 7, Appli

29	48.8	3.0	8854	3	US-09-053-549-1 Sequence 1, Appli
30	48.4	2.9	1531	1	US-08-449-986-1 Sequence 1, Appli
31	48.4	2.9	1531	2	US-08-756-855-1 Sequence 1, Appli
32	48.2	2.9	1187	1	US-08-440-856A-2 Sequence 2, Appli
33	48.2	2.9	38506	3	US-09-320-878-19 Sequence 19, Appli
34	48.2	2.9	44377	2	US-08-804-227C-7 Sequence 7, Appli
35	48.2	2.9	44377	2	US-08-804-198-1 Sequence 1, Appli
36	47.8	2.9	1052	2	US-08-403-852D-10 Sequence 10, Appli
37	47.8	2.9	1052	3	US-08-510-646B-10 Sequence 10, Appli
38	47.4	2.9	1288	1	US-08-440-856A-9 Sequence 9, Appli
39	46.8	2.9	2943	1	US-08-042-747A-7 Sequence 7, Appli
40	46.6	2.8	2064	1	US-08-343-428-1 Sequence 1, Appli
41	46.2	2.8	1389	2	US-08-023-980B-3 Sequence 3, Appli
42	46.2	2.8	1389	2	US-08-486-953A-3 Sequence 3, Appli
43	45.6	2.8	774	3	US-08-956-307B-12 Sequence 12, Appli
44	45.6	2.8	778	3	US-08-956-307B-11 Sequence 11, Appli
45	45.2	2.8	1931	2	US-09-130-114-2 Sequence 2, Appli

#### ALIGNMENTS

RESULT 1  
US-07-951-715A-6  
Sequence 6, Application US/07951715A  
Patent No. 5625136  
GENERAL INFORMATION:  
APPLICANT: Kozziel, Michael G.  
APPLICANT: Desai, Nalini M.  
APPLICANT: Lewis, Kelly S.  
APPLICANT: Kramer, Vance C.  
APPLICANT: Warren, Gregory W.  
APPLICANT: Ewola, Stephen V.  
APPLICANT: Crossland, Lyle D.  
APPLICANT: Wright, Martha S.  
APPLICANT: Merlin, Ellis J.  
APPLICANT: Launus, Karen L.  
APPLICANT: Rothstein, Steven J.  
APPLICANT: Bowman, Cindy G.  
APPLICANT: Dawson, John L.  
APPLICANT: Dunder, Erik M.  
APPLICANT: Pace, Gary M.  
APPLICANT: Suttie, Janet L.  
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED  
TITLE OF INVENTION: INSECTICIDAL ACTIVITY IN MAIZE  
NUMBER OF SEQUENCES: 94  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30B  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/951,715A  
FILING DATE: 25-SEP-1992  
CLASSIFICATION: 800  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 07/772,027  
FILING DATE: 04-OCT-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Spruill, W. Murray  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: S-18805/A/CGC 1577/CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8615  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 6:

```

?      SEQUENCE CHARACTERISTICS:
?      LENGTH: 3624 base pairs
?      TYPE: nucleic acid
?      STRANDEDNESS: single
?      TOPOLOGY: linear
?      MOLECULE TYPE: other nucleic acid
?      DESCRIPTION: /desc = "Synthetic DNA"
?      HYPOTHEetical: NO
?      FEATURE:
?      NAME/KEY: CDS
?      LOCATION: 1..3621
?      OTHER INFORMATION:
?      OTHER INFORMATION:
?      OTHER INFORMATION:
?      /note= "Disclosed in Figure 6."
US-07-951-715A-6

```

[illegible]

RESULT 2  
US-08-459-448A-6  
Sequence 6, Application US/08459448A  
Patent No. 5859336  
GENERAL INFORMATION:  
APPLICANT: Koziel, Michael G.  
APPLICANT: Desai, Nalini M.  
APPLICANT: Lewis, Kelly S.  
APPLICANT: Kramer, Vance C.  
APPLICANT: Warren, Gregory W.  
APPLICANT: Evola, Stephen V.  
APPLICANT: Crossland, Lyle D.  
APPLICANT: Wright, Martha S.  
APPLICANT: Merlin, Ellis J.  
APPLICANT: Launis, Karen L.  
APPLICANT: Rothstein, Steven J.  
APPLICANT: Bowman, Cindy G.  
APPLICANT: Dawson, John L.  
APPLICANT: Dunder, Erik M.  
APPLICANT: Pace, Gary M.  
APPLICANT: Suttie, Janet L.  
TITLE OF INVENTION: SYNTHETIC DNA SEQUENCE HAVING ENHANCED  
INSECTICIDAL ACTIVITY IN MAIZE

```

1  NUMBER OF SEQUENCES: 94
2  CORRESPONDENCE ADDRESS:
3  ADDRESSEE: No. 585933Cartis Corporation
4  STREET: Patent & Trademark Dept., 520 White Plains
5  STREET: Rd., POB 2005
6  CITY: Tarrytown
7  STATE: New York
8  COUNTRY: USA
9  ZIP: 10591-9005
10
11  COMPUTER READABLE FORM:
12  MEDIUM TYPE: Floppy disk
13  COMPUTER: IBM PC compatible
14  OPERATING SYSTEM: PC-DOS/MS-DOS
15  SOFTWARE: PatentIn Release #1.0, Version #1.30
16  CURRENT APPLICATION DATA:
17  APPLICATION NUMBER: US/08/459,448A
18  FILING DATE: 02-JUN-1995
19  CLASSIFICATION: 800
20  PRIOR APPLICATION DATA:
21  APPLICATION NUMBER: US 07/951,715
22  FILING DATE: 25-SEP-1992
23  PRIOR APPLICATION DATA:
24  APPLICATION NUMBER: US 07/772,027
25  FILING DATE: 04-OCT-1991
26  ATTORNEY/AGENT INFORMATION:
27  NAME: Pace, Gary M.
28  REGISTRATION NUMBER: 40403
29  REFERENCE/DOCKET NUMBER: CGC 1577/CIP/DIV4
30  TELECOMMUNICATION INFORMATION:
31  TELEPHONE: (919)541-8582
32  TELEFAX: (919)541-8689
33  INFORMATION FOR SEQ ID NO: 6:
34  SEQUENCE CHARACTERISTICS:
35  LENGTH: 3624 base pairs
36  TYPE: nucleic acid
37  STRANDEDNESS: single
38  TOPOLOGY: linear
39  MOLECULE TYPE: other nucleic acid
40  DESCRIPTION: /desc = "Synthetic DNA"
41  HYPOTHETICAL: NO
42  FEATURE:
43  NAME/KEY: CDS
44  LOCATION: 1..3621
45  OTHER INFORMATION:
46  OTHER INFORMATION: /product= "Full-length, maize
47  OTHER INFORMATION: optimized cryiB"
48  OTHER INFORMATION: /note= "Disclosed in Figure 6."
49
50  /S-08-459-448A-6

```

	Query Match	3.4%;	Score 56;	DB 2;	Length 3624;
	Best Local Similarity	48.4%;	Pred. No. 0.0008;		
	Matches 186;	Conservative 0;	Mismatches 195;	Indels 3;	Gaps 1;
QY	83 gactgtccccgcaggggacaggcccgccagaccgcgcgtcttcgscgcgccaccaacctgcg	142			
Db	2286 GATGGGCGAGAGCCAGCTGAAGGCTTAACCCGGCTAACAGCTGCCTTACATGAGAGA	2345			
QY	143 cagcgtgtgcgccttgygcctaacaatctctacgaccgcggcgagtatacccagttacct	202			
Db	2346 CAGCAGGACTGTGAGANTCACTGATCGGTGTACAAGCCCAAAGCATCAGACCCTTGAGCT	2405			
QY	203 gcctgaagacctactgtccttgtgcgggggctcgtcaaccggtgttctgcgcgaagaagacgt	262			
Db	2406 GCCCGGACCCAGAGACTGTGTGCCCTTCGAGCGTAGAGGCCCAATCGGCGGTGGGGGA	2465			
QY	263 ggcgttcgcagcgcgcctctgtctacatgaccacgtgtccttgcgcgcaccccgctg	322			
Db	2466 GCCCACCGCTGCGCCCCCCCACCTTCGAGTTGAGAACCCCGACCTGAGACTCACACTGCC---G	2522			
QY	323 cgccgcgggcgcgttcgtcttaccgcggagccttaagtaaccattccccgtygggtcgaactg	382			
Db	2523 CGACGCGGAGAAATGCGCCCAACCAAGCACACCACTTCAAGCTTGACATCGAGCTGGCGTTG	2582			







3. Inn  
D0 2153 agctgaccgtgcccaaggagggcccccgtgaggggtccctgagg 2195

RESULT 9 Application US/09135021A  
US-09-135-021-5  
Sequence 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836,

```

NAME/KEY: CUD..(1011)
LOCATION: (153)..(1011)
FEATURE: mutation
NAME/KEY: (730)
LOCATION/EN: (730)
OTHER INFORMATION: wild-type.
OTHER INFORMATION:
US-03-135-021-5
3.3%; Score 53.4; DB 3; length 3182
At 0.0.0031;bi; indels

```

[illegible]

DB

	XY	436	agggcgccgaattt	TTTTT
		TT	TT	TT
	2154	agctaccggtccacagaaggcccccatr		

RESULT 10  
US-08-471-601-23 Application US C.C. Nuclear Genetic System For  
Sequence 23, Applicant:  
Patent No. 5689049  
PATENT INFORMATION: CITAN  
GENERIC APPLICANT: ALBERT  
APPLICANT:  
TITLE OF INVENT

APPLICANT: ALBERTSEN, Marc C.  
 TITLE OF INVENTION: Reversible Nuclear Genetic System For  
 TITLE OF INVENTION: Male Sterility In Transgenic Plants  
 NUMBER OF SEQUENCES: 23  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Foley & Lardner  
 STREET: 3000 K Street, N.W., Suite 500  
 CITY: Washington  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20007-5109  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/474,556  
 FILING DATE: 07-JUN-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/351,899  
 FILING DATE: 08-DEC-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: BENT, Stephen A.  
 REGISTRATION NUMBER: 29,768  
 REFERENCE/DOCKET NUMBER: 33229/329/PIHI  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (202)672-5300  
 TELEFAX: (202)672-5399  
 TELEX: 904136  
 INFORMATION FOR SEQ ID NO: 23:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1485 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear

Query Match 1: 3.28; Score 52.6; DB 1; Length 1485;

	Matches	164;	Conservative	0;	Mismatches	164;	Indels	3;	Gaps	1
OY	71	g g c g c g c a a g c g n a g t c g c c c g g c a a g g g c a a g c c c g c a a c g c g t t c c g g c g c d								130
Db	486	g t c g a g c c a g a g g c g c c t t c c c g g g c g c a c c t g a c a c t g g g c g g c c c t g g g c t g a d								545
OY	131	c a c c a c a c t c g c a g c g t g c g c c c t g g g c t a c a g a t c t c t a c a g a c c g g a g a g t a								190
Db	546	c c c g g a c g t c g c g c c g c g a t c t g g a c c t t t a c c g g c t c g t c g g g c g g g t g c g g g c t								605
OY	191	c c c c a g t a c c t c g t a a g c t a --- c t g c t g t g c g g g t c g t c t a a c c g a g t g t g t								247
Db	606	c c g c t g g c c a a g g c c t g e c c g a g a c t g c c c g g g c g c g c g c g t c t g c t g c c a c t t c								665
OY	248	c g g c g a g a g a g a g c g t c g c t c c g a g c g c c c t g t c t a t g a t g c c a c g t c g t c t c g								307
Db	666	c g a g a c c a c a c a t g t g g g g t t c c g c c c g c c a g c c c g a c c g c c c t a c g a c c t c c t g g g								725
OY	308	c c g a c c c c c g g c t c t c c g c g c g c g c g t t c g t c t a a c a c a g a g c t a g t a c a c t c c								367
Db	726	c g t g g c g c t c t t t c g g c g a g a c c g c c g g g c c c c g c c t g a t c a t c g g a c c a c c c c c c c c								785
OY	368	c g t g g c t g a c c t g c t c c c c g a a c c g a g								398

RESULT 12  
US-08-479-382-23  
; Sequence 23. Application US/08479382  
; Patent No. 5763243







Db 786 CGAGCGCCCGCTCTTCGAGCTCCACTCGCG 816

Search completed: February 24, 2001, 21:06:35  
Job time: 6993 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2001, 21:51:30 ; Search time 119.29 Seconds

(without alignments)  
5170.912 Million cell updates/sec

Title: US-09-320-713-3

Perfect score: 1642  
Sequence: 1 ggaattcgccgcgcgcgtcgt.....tgaaaaaaaaaaaaaaaaaa 1642

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_36.\*  
1: /SID56/gcgdata/geneseq/geneseqn/NA1980.DAT.\*  
2: /SID56/gcgdata/geneseq/geneseqn/NA1981.DAT.\*  
3: /SID56/gcgdata/geneseq/geneseqn/NA1982.DAT.\*  
4: /SID56/gcgdata/geneseq/geneseqn/NA1983.DAT.\*  
5: /SID56/gcgdata/geneseq/geneseqn/NA1984.DAT.\*  
6: /SID56/gcgdata/geneseq/geneseqn/NA1985.DAT.\*  
7: /SID56/gcgdata/geneseq/geneseqn/NA1986.DAT.\*  
8: /SID56/gcgdata/geneseq/geneseqn/NA1987.DAT.\*  
9: /SID56/gcgdata/geneseq/geneseqn/NA1988.DAT.\*  
10: /SID56/gcgdata/geneseq/geneseqn/NA1989.DAT.\*  
11: /SID56/gcgdata/geneseq/geneseqn/NA1990.DAT.\*  
12: /SID56/gcgdata/geneseq/geneseqn/NA1991.DAT.\*  
13: /SID56/gcgdata/geneseq/geneseqn/NA1992.DAT.\*  
14: /SID56/gcgdata/geneseq/geneseqn/NA1993.DAT.\*  
15: /SID56/gcgdata/geneseq/geneseqn/NA1994.DAT.\*  
16: /SID56/gcgdata/geneseq/geneseqn/NA1995.DAT.\*  
17: /SID56/gcgdata/geneseq/geneseqn/NA1996.DAT.\*  
18: /SID56/gcgdata/geneseq/geneseqn/NA1997.DAT.\*  
19: /SID56/gcgdata/geneseq/geneseqn/NA1998.DAT.\*  
20: /SID56/gcgdata/geneseq/geneseqn/NA1999.DAT.\*  
21: /SID56/gcgdata/geneseq/geneseqn/NA2000.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1642	100.0	1642	21	236835	Partial nucleotide
2	1615.6	98.4	2361	21	252201	Human transforming
3	1598.6	97.4	1819	21	252195	Human transforming
4	463.4	28.2	522	21	236837	Partial nucleotide
5	326.4	19.9	1221	21	252198	Murine transformin
6	239.6	14.6	498	21	236853	CDNA clone H4CBX04
7	196	11.9	371	21	236852	CDNA clone H4CBX04
8	167.4	10.2	178	21	236854	CDNA clone H4CBX04
9	165.4	10.1	264	21	236855	CDNA clone H4CBX04
10	73.2	4.5	540	20	H89401	Human interleukin-10
11	73.2	4.5	687	20	234282	Human interleukin-10
12	73.2	4.5	687	21	229727	Human interleukin-10

13	73.2	4.5	705	20	X15870	CDNA encoding inte
14	73.2	4.5	736	20	V83172	Human zcyto7 CDNA
15	73.2	4.5	738	20	X84472	Human EDIRP I cod
16	73.2	4.5	1259	20	X84474	Human EDIRP II cod
17	71.8	4.4	100	16	T23250	Human gene signatu
18	61.6	3.8	114955	20	X53491	Human adenosine A1
19	60.2	3.7	114955	20	X53491	Human adenosine A1
20	59.4	3.6	692	20	V83179	Murine zcyto7 CDNA
21	59.4	3.6	736	20	X84473	Murine EDIRP I cod
22	58.8	3.6	540	20	X89402	Mouse interleukin-
23	58.8	3.5	591	21	A09153	Human interleukin-
24	56.8	3.5	1047	21	Z29728	Human interleukin
25	56.8	3.5	1067	21	Z36836	Nucleotide sequenc
26	56.8	3.5	4356	16	O95540	Cardiac adenylyl c
27	56.2	3.4	1145	16	T04785	Cellulose gene fr
28	55.2	3.4	2291	9	N80309	Entire amyase gene
29	55.2	3.4	4356	14	O37543	Cardiac adenylyl c
30	53.6	3.3	265	20	X84482	Probe for murine E
31	53.6	3.3	497	20	V83181	Murine zcyto7 expr
32	53.6	3.3	1127	21	A02477	Human colon cancer
33	53.6	3.3	2271	11	O06844	Amylase gene from
34	53.4	3.3	2821	18	T94004	DNA encoding human
35	53.4	3.3	2821	18	T90730	Human KVLQ1 full-
36	53.4	3.3	3181	21	Z90669	Human KVLQ1 prote
37	53.4	3.3	3181	21	Z98901	Human long QT synd
38	53	3.2	1059	21	A01833	Human colon cancer
39	53	3.2	24379	19	T93095	Streptomyces iteno
40	52.6	3.2	24379	19	V25925	Streptomyces roseo
41	52.6	3.2	1485	17	T35134	5126 another specif
42	52.4	3.2	830	21	Z29729	Virtual DNA fragme
43	52	3.2	3198	20	X02974	Human IL-1ra BAC c
44	51.6	3.1	1497	11	O03367	Phenol oxidase (PO
45	51.6	3.1	1497	11	O03371	Sequence encoding

#### ALIGNMENTS

RESULT 1	
ID 236835	236835 standard; cDNA; 1642 BP.
XX	
AC 236835;	
XX	
DT 13-MAR-2000 (first entry)	
DE	Partial nucleotide sequence of human interleukin-22.
XX	
XX	Human: interleukin-22; IL-22; IL-21; immune system disorder;
KW	immune cell chemotaxis; haematopoietic cell disorder;
KW	haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;
KW	respiratory problem; organ rejection; graft-versus-host disease; GVHD;
KW	inflammation; hyperproliferative disorder; tissue regeneration;
KW	embryonic stem cell differentiation; embryonic stem cell proliferation;
KW	haematopoietic lineage; allergic asthma; ss.
XX	
OS Homo sapiens.	
XX	
EH	
FT CDS	Location/Qualifiers
FT	3..485
FT	/product= "partial IL-22"
PD	
XX	
XX	02-DEC-1999.
XX	
XX	27-MAY-1999; 99MO-US11644.
XX	
XX	29-MAY-1998; 98US-0087340.
PR	10-SEP-1998; 98US-009805.
PR	30-APR-1999; 99US-0131965.
XX	

PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Ruben SM, Edner R;  
XX  
DR WPI: 2000-072622/06.  
DR P-PSDB: Y53891.  
XX  
PT Novel polynucleotides used to develop products for treating e.g. immune  
PT disorders, blood disorders, autoimmune disorders, allergies,  
PT inflammation, hyperproliferative disorders or infections  
XX  
PS Claim 12; Fig 2A-B; 170pp; English.  
XX  
CC The present sequence encodes a partial human interleukin-22 (IL-22)  
CC protein. The specification also describes IL-21 polynucleotides and  
CC polypeptides. The IL-21 polynucleotide was isolated from a cDNA library  
CC of epileptic frontal cortex. IL-21 and IL-22 may be useful in treating  
CC deficiencies or disorders of the immune system, by activating or  
CC inhibiting the proliferation, differentiation, or mobilization  
CC (chemotaxis) of immune cells, treating or detecting deficiencies  
CC of haematopoietic cells, to modulate haemostatic or thrombolytic  
CC activity, in treating or detecting autoimmune disorders, treating  
CC asthma (particularly allergic asthma) or other respiratory problems,  
CC to treat and/or prevent organ rejection or graft-versus-host disease  
CC (GVHD), to modulate inflammation, to treat or detect hyperproliferative  
CC disorders, to treat or detect infectious agents, to differentiate,  
CC proliferate and attract cells, leading to the regeneration of tissues,  
CC IL-21 and IL-22 may also increase or decrease the differentiation or  
CC proliferation of embryonic stem cells and haematopoietic lineage, may  
CC be used to modulate mammalian characteristics.  
XX  
XX Sequence 1642 BP; 377 A; 458 C; 439 G; 368 T; 0 other:  
50  
Query Match 100.0%; Score 1642; DB 21; Length 1642;  
Best Local Similarity 100.0%; Pred. NO. 0;  
Matches 1642; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ggaattggcagcagactgtgtcgcgtgcagtcagcttcacacacagctgacgtgggc 60  
DB 1 ggaattggcagcagactgtgtcgcgtgcagtcagcttcacacacagctgacgtgggc 60  
QY 61 cgcgtgagcagc 120  
DB 61 cgcgtgagcagc 120  
QY 121 tccgc 180  
DB 121 tccgc 180  
QY 181 cggcgaaggtaccccaagctgacgtgcgtgacgtgcgtgcgtgcgtgcgtgcgtgc 240  
DB 181 cggcgaaggtaccccaagctgacgtgcgtgacgtgcgtgcgtgcgtgcgtgcgtgc 240  
QY 241 ggcgtgttcgc 300  
DB 241 ggcgtgttcgc 300  
QY 301 tccctgc 360  
DB 301 tccctgc 360  
QY 361 ccaatcccggtggtgc 420  
DB 361 ccaatcccggtggtgc 420  
QY 421 ccaagatcgcacacacagc 480  
DB 421 ccaagatcgcacacacagc 480  
QY 481 cctgagcgcgcgttcctgc 540  
DB 481 cctgagcgcgcgttcctgc 540

QY 541 gaggcgcgtggaggc 600  
DB 541 gaggcgcgtggaggc 600  
QY 601 ccgagac 660  
DB 601 ccgagac 660  
QY 661 tccctgc 720  
DB 661 tccctgc 720  
QY 721 cggcgaaggtaccccaagctgacgtgcgtgacgtgcgtgacgtgcgtgacgtgcgt 780  
DB 721 cggcgaaggtaccccaagctgacgtgcgtgacgtgcgtgacgtgcgtgacgtgcgt 780  
QY 781 ctgaggaagccttcac 840  
DB 781 ctgaggaagccttcac 840  
QY 841 gccaaagagataggagacgcacatgcttttaagcaatcctaaataataagtagtag 900  
DB 841 gccaaagagataggagacgcacatgcttttaagcaatcctaaataataagtagtag 900  
QY 901 cgaactatcactccttttaaacatcactgttttgatagagcagcactatctatatt 960  
DB 901 cgaactatcactccttttaaacatcactgttttgatagagcagcactatctatatt 960  
QY 961 atcaaataggcactccttcgttaacattcttcaacataataacacgcgttttactcttc 1020  
DB 961 atcaaataggcactccttcgttaacattcttcaacataataacacgcgttttactcttc 1020  
QY 1021 tggtagaatttttaagaacataattggaaatccctggataaattttgtagctgtacac 1080  
DB 1021 tggtagaatttttaagaacataattggaaatccctggataaattttgtagctgtacac 1080  
QY 1081 tggcctgtgtcttcgaattcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1140  
DB 1081 tggcctgtgtcttcgaattcagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1140  
QY 1141 catctgaccacactccttccttcacatgaaaggtcttcacgcgcgcgcgcgcgcgcgcgc 1200  
DB 1141 catctgaccacactccttccttcacatgaaaggtcttcacgcgcgcgcgcgcgcgcgcgc 1200  
QY 1201 atgcacagc 1260  
DB 1201 atgcacagc 1260  
QY 1261 ttttagtcatgatacaataaacaagctcacaactgcacaatttttccctcttgaag 1320  
DB 1261 ttttagtcatgatacaataaacaagctcacaactgcacaatttttccctcttgaag 1320  
QY 1321 ccaactgaggccaatttgcgttaagaagtggtgagataaagaagtggaaagtcgacatctt 1380  
DB 1321 ccaactgaggccaatttgcgttaagaagtggtgagataaagaagtggaaagtcgacatctt 1380  
QY 1381 gccagttgtcagaagaatcccaagcagatgagcttgaaggtcttaagatacagg 1440  
DB 1381 gccagttgtcagaagaatcccaagcagatgagcttgaaggtcttaagatacagg 1440  
QY 1441 ctgaaatgaggaacaagaatgggcacgcttaagcattgaaagatacaactctgaagcttct 1500  
DB 1441 ctgaaatgaggaacaagaatgggcacgcttaagcattgaaagatacaactctgaagcttct 1500  
QY 1501 gttctgtcattctgcacagaaagctaggctccttgatccttcttaagattgaagctgt 1560  
DB 1501 gttctgtcattctgcacagaaagctaggctccttgatccttcttaagattgaagctgt 1560  
QY 1561 ctctgaacaacaattattgtataaagttagtagttctttttaaatcatataaagaagctt 1620  
DB 1561 ctctgaacaacaattattgtataaagttagtagttctttttaaatcatataaagaagctt 1620

OY 1621 gctgaataaaaaaa 1642  
|||||  
Db 1621 gctgaataaaaaaa 1642

RESULT 2  
ID 252201  
252201 standard; cDNA: 2361 BP.

XX 252201;

DT 18-JUL-2000 (first entry)

DE Human transforming growth factor beta-9, Ztgbeta-9 variant cDNA.

KM Human transforming growth factor beta-9; Ztgbeta-9;  
KM Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
KM amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
KM peripheral neuropathy; demyelinating disease; multiple sclerosis;  
KM antiviral; cyostatic; ss.

XX Homo sapiens.

OS Homo sapiens.  
FH Key Location/Qualifiers  
FT CDS 572..1201  
FT /tag= a  
FT /product= "Human Ztgbeta-9 protein"

FT sig\_peptide 76..637

FT mat\_peptide 638..1198

FT /tag= b  
FT /product= "Mature Ztgbeta-9 protein"

XX WO200015798-A2.

PN 23-MAR-2000.

PD 17-SEP-1999; 99WO-US21677.

PF 17-SEP-1998; 98US-0154817.

PR (ZYMO) ZYMOGENETICS INC.

PA Presnell SR, Taft DW, Foley KP;

PI WPI: 2000-271436/23.

DR P-PSDB; Y70662, Y70663.

XX Polynucleotides encoding a novel transforming growth factor beta-9  
PT polypeptide, designated Ztgbeta-9, useful as an antiviral and  
PT antiproliferative agent.

XX Example 1; Page 90-92; 97pp; English.

CC The present sequence encodes a variant of human transforming growth  
CC factor beta-9, designated Ztgbeta-9. Human Ztgbeta-9 was isolated  
CC from an arrayed pituitary gland cDNA plasmid library by PCR screening.  
CC This can be used to treat a variety of neurodegenerative diseases such  
CC as amyotrophic lateral sclerosis (ALS), Alzheimer's disease, Huntington's  
CC disease, Parkinson's disease and peripheral neuropathies, or  
CC demyelinating diseases including multiple sclerosis. Ztgbeta-9 peptides  
CC have antiviral activity and may also be used to regulate the  
CC proliferation, differentiation and apoptosis of neurons, glial cells,  
CC lymphocytes, hematopoietic cells and stromal cells.

CC Sequence 2361 BP; 509 A; 687 C; 655 G; 510 T; 0 other;

Query Match 98.4%; Score 1615.6; DB 21; Length 2361;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1618; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 21 gccgtgctcagtcctccacacacgcgtcagctggcgccgtgtagcagcgccgaac 80

Db 737 ggcgtgctcagtcctccacacacgcgtcagctggcgccgtgtagcagcgccgaac 796  
OY 81 gcgagctgcccggcgaaggcgaagcccgccagccgcttccggccgccacaactctg 140  
Db 797 gcgagctgcccggcgaaggcgaagcccgccagccgcttccggccgccacaactctg 856  
OY 141 cgcagcgtgctgcccggcgaagctcagatctctacagcccggaaggtaccaggtac 200  
Db 857 cgcagcgtgctgcccggcgaagctcagatctctacagcccggaaggtaccaggtac 916  
OY 201 ctgctgaagcctactgctcgttccggtgctcgtcagccgggtcttccggaaggagac 260  
Db 917 ctgctgaagcctactgctcgttccggtgctcgtcagccgggtcttccggaaggagac 976  
OY 261 gtgcgttccggaagcctcgttctacatgcccacgctgctctgcgcgcaccccgcc 320  
Db 977 gtgcgttccggaagcctcgttctacatgcccacgctgctctgcgcgcaccccgcc 1036  
OY 321 tgc 380  
Db 1037 tgc 1096  
OY 381 tgc 440  
Db 1097 tgc 1156  
OY 441 gccaaagctcctgtctgcccacacgacgcgcgcgcgcgcgcgcgcgcgcgcgcgc 500  
Db 1157 gccaaagctcctgtctgcccacacgacgcgcgcgcgcgcgcgcgcgcgcgcgcgc 1216  
OY 501 gggaggtctcccccgc 560  
Db 1217 gggaggtctcccccgc 1276  
OY 561 tccggaagcctctgaagagctgacacggaacacgaagtcggaagcgcgcgcgcgcgc 620  
Db 1277 tccggaagcctctgaagagctgacacggaacacgaagtcggaagcgcgcgcgcgcgc 1336  
OY 621 tccatgagagcctctgaagagctgacacggaacacgaagtcggaagcgcgcgcgcgcgc 680  
Db 1337 tccatgagagcctctgaagagctgacacggaacacgaagtcggaagcgcgcgcgcgcgc 1396  
OY 681 acaagcaagcagcgtgctggaagctgataagtcgaagcgcgcgcgcgcgcgcgcgcgcgc 740  
Db 1397 acaagcaagcagcgtgctggaagctgataagtcgaagcgcgcgcgcgcgcgcgcgcgcgc 1456  
OY 741 cggccgc 800  
Db 1457 cggccgc 1516  
OY 801 gctgctgcccggcgaaggc 860  
Db 1517 gctgctgcccggcgaaggc 1576  
OY 861 tatgcttttaagaacatctaaataataataatagatagagctatataccacttta 920  
Db 1577 tatgcttttaagaacatctaaataataataatagatagagctatataccacttta 1636  
OY 921 aatcaacgttttgaataagcgaagcatttatatataatagaagcattctg 980  
Db 1637 aatcaacgttttgaataagcgaagcatttatatataatagaagcattctg 1696  
OY 981 ttacatttcttaacataaacaatcgttttctacttcttctgtagaatttttaagca 1040  
Db 1697 ttacatttcttaacataaacaatcgttttctacttcttctgtagaatttttaagca 1756  
OY 1041 taattggaatccttgataaattttgtagctgttaaacctctggtctgtaattc 1100  
Db 1757 taattggaatccttgataaattttgtagctgttaaacctctggtctgtaattc 1816  
OY 1101 agcctgaacgagtgctgctgataaaggaagcagctccactgaacacacttccct 1160

Db 1817 agcgtaccgatggtcgtactgataatgacacgtctcatctgacccactcttcctt 1876  
Qy 1161 ccactgaaggtcttcacggcgtccacgttgacccaagaagatgacagggcgctgcacgtg 1220  
Db 1877 ccactgaaggtcttcacggcgtccacgttgacccaagaagatgacagggcgctgcacgtg 1936  
Qy 1221 ccccaaggcgcagctaaagatgctccaaagatctccagattggttttagtcaataacataa 1280  
Db 1937 ccccaaggcgcagctaaagatgctccaaagatctccagattggttttagtcaataacataa 1996  
Qy 1281 acagctcaaacctgcacacatttttcccccttttgaagccactggggccaatttgtg 1340  
Db 1997 acagctcaaacctgcacacatttttcccccttttgaagccactggggccaatttgtg 2056  
Qy 1341 ttaagagggtgtgagataagaagtggaaagcttgcacattcttcacagttgcagaagaacc 1400  
Db 2057 ttaagagggtgtgagataagaagtggaaagcttgcacattcttcacagttgcagaagaacc 2116  
Qy 1401 aagcaggtatggtcttaaggttgagaggtcttagatcacagctgaataatgagacaagctg 1460  
Db 2117 aagcaggtatggtcttaaggttgagaggtcttagatcacagctgaataatgagacaagctg 2176  
Qy 1461 ggcacagcttagcatctgcagagatcaatctgagaggtctgtttctgcattctgcacga 1520  
Db 2177 ggcacagcttagcatctgcagagatcaatctgagaggtctgtttctgcattctgcacga 2236  
Qy 1521 gagctaggtcttgaatcttctttagattgaagctgctctcgaacaacattatttgt 1580  
Db 2237 gagctaggtcttgaatcttctttagattgaagctgctctcgaacaacattatttgt 2296  
Qy 1581 aaaagttagttcttcttttaaatcatataaagaggtctgtgaaaaaataaaaaa 1640  
Db 2297 aaaagttagttcttcttttaaatcatataaagaggtctgtgaaaaaataaaaaa 2356  
Qy 1641 aa 1642  
Db 2357 aa 2358

RESULT 3  
252195  
ID 252195 standard; cDNA; 1819 BP.  
XX  
AC 252195;  
Df 18-JUL-2000 (first entry)  
XX  
De Human transforming growth factor beta-9, Ztgifbeta-9 cDNA.  
XX  
Xx Human transforming growth factor beta-9; Ztgif beta-9;  
KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
KW antiviral; cytosolic; ss.  
OS  
XX Homo sapiens.  
XX  
FH Key location/Qualifiers  
FT CDS 71..679  
FT /tag= a  
FT /product= "Ztgifbeta-9 protein"  
FT s1g\_peptide 1..115  
FT /tag= b  
FT mat\_peptide 116..676  
FT /tag= c  
FT /product= "Mature Ztgifbeta-9 protein"  
FT mat\_peptide 80..676  
FT /tag= d  
FT /product= "Mature protein represented as SEQ ID NO:3"  
FT mat\_peptide 83..676  
FT /tag= e  
FT /product= "Mature protein represented as SEQ ID NO:4"  
FT mat\_peptide 86..676

FT /tag= f  
FT /product= "Mature protein represented as SEQ ID NO:5"  
XX  
PN WO200015798-A2.  
XX  
PD 23-MAR-2000.  
XX  
PF 17-SEP-1999; 99WO-US21677.  
XX  
PR 17-SEP-1998; 98US-0154817.  
XX  
XX (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Presnell SR, Taft DW, Foley KP;  
XX  
DR WPI: 2000-271436/23.  
DR P-PSDB: Y70653, Y70654, Y70655, Y70656.  
XX  
PT Polynucleotides encoding a novel transforming growth factor beta-9  
PT polypeptide, designated Ztgif beta-9, useful as an antiviral and  
PT antiproliferative agent -  
XX  
PS Claim 5; Page 81-83; 97pp; English.  
XX  
CC The present sequence encodes human transforming growth factor beta-9,  
CC designated Ztgif beta-9. Human Ztgif beta-9 was isolated from an arrayed  
CC pituitary gland cDNA plasmid library by PCR screening. This can be used  
CC to treat a variety of neurodegenerative diseases such as amyotrophic  
CC lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,  
CC Parkinson's disease and peripheral neuropathies, or demyelinating  
CC diseases including multiple sclerosis. Ztgif beta-9 peptides have  
CC antiviral activity and may also be used to regulate the proliferation,  
CC differentiation and apoptosis of neurons, glial cells, lymphocytes,  
CC hematopoietic cells and stromal cells.  
XX  
SQ Sequence 1819 BP; 374 A; 528 C; 530 G; 387 T; 0 other:  
  
Query Match 97.4%; Score 1598.6; DB 21; Length 1819;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1601; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
  
Qy 21 ggcgtgctcagtgctcttcacacacacagctgagctgggcccgtgagcagcgcgcaac 80  
Db 215 ggcgtgctcagtgctcttcacacacacagctgagctgggcccgtgagcagcgcgcaac 274  
Qy 81 ggcagctgcccgcagaggggagggccgcgcagccgcgtctccggcccccacaacctg 140  
Db 275 ggcagctgcccgcagaggggagggccgcgcagccgcgtctccggcccccacaacctg 334  
Qy 141 gcagcgtgtgccttggtgcttaccagaatctctaccagcccgagaggtaccagggtac 200  
Db 335 gcagcgtgtgccttggtgcttaccagaatctctaccagcccgagaggtaccagggtac 394  
Qy 201 ctgctgaagcctaactgctgtgcccgggctgcttgcagcgggctgttgcggagagagac 260  
Db 395 ctgctgaagcctaactgctgtgcccgggctgcttgcagcgggctgttgcggagagagac 454  
Qy 261 gtgcgtctccgagagccctgttctaccatgccaacgctgctcgtgcgcgcaccccgcc 320  
Db 395 ctgctgaagcctaactgctgtgcccgggctgcttgcagcgggctgttgcggagagagac 454  
Qy 455 gtgcgtctccgagagccctgttctaccatgccaacgctgctcgtgcgcgcaccccgcc 514  
Db 321 tgcgcgcgcgcgcgtcttccgttaccacgagcctaccagtcacacaccccggtggtgac 380  
Qy 515 tgcgcgcgcgcgcgtcttccgttaccacgagcctaccagtcacacaccccggtggtgac 574  
Db 381 tgcgtcccccagccggaagagagcagcagcaltcaactcagcagcatcgacaacagggc 440  
Qy 575 tgcgtcccccagccggaagagagcagcagcaltcaactcagcagcatcgacaacagggc 634  
Db 441 gccaaagctcctgtgtggcccacaagcagcgcgcgtgcccctgaagcggctcctgcc 500  
Qy 635 gccaaagctcctgtgtggcccacaagcagcgcgcgtgcccctgaagcggctcctgcc 694

QY 501 gggaggtctcccgcccgcatcccgagggcccaagctggagccgctggagggctcgg 560  
 |||||  
 Db 695 gggaggtctcccgcccgcatcccgagggcccaagctggagccgctggagggctcgg 754  
 QY 561 tcggcgacctctgaagagatgcacccagcaaaccaagctgcggagacacagcgccgct 620  
 |||||  
 Db 755 tcggcgacctctgaagagatgcacccagcaaaccaagctgcggagacacagcgccgct 814  
 QY 621 ttccatggagacctcgttaagagcttcctgcacaacgggacccctgctgttttagct 680  
 |||||  
 Db 815 ttccatggagacctcgttaagagcttcctgcacaacgggacccctgctgttttagct 874  
 QY 681 acaagcaagcagcgttgctggaagctatgtggaacgaacccggcacgggacccctgtgt 740  
 |||||  
 Db 875 acaagcaagcagcgttgctggaagctatgtggaacgaacccggcacgggacccctgtgt 934  
 QY 741 cggcccgcatggaaggttttgaaaagttcacaggaggtcccttgaggagaccttcagatcg 800  
 |||||  
 Db 935 cggcccgcatggaaggttttgaaaagttcacaggaggtcccttgaggagaccttcagatcg 994  
 QY 801 gctgctcgcggtgcaaggcgctgaactacgcgtggtgcttgcacaaagagataaggagcga 860  
 |||||  
 Db 995 gctgctcgcggtgcaaggcgctgaactacgcgtggtgcttgcacaaagagataaggagcga 1054  
 QY 861 tatgcttttaagacatctaaataataataagatagcagctatatacctactttta 920  
 |||||  
 Db 1055 tatgcttttaagacatctaaataataataagatagcagctatatacctactttta 1114  
 QY 921 aatacaactgttttgaatagagcagagcattttatatacaaatgaaagactactctg 980  
 |||||  
 Db 1115 aatacaactgttttgaatagagcagagcattttatatacaaatgaaagactactctg 1174  
 QY 981 ttacatttcttaacataataaacatcglttttcttctctctgtagaatltttaagca 1040  
 |||||  
 Db 1175 ttacatttcttaacataataaacatcglttttcttctctctgtagaatltttaagca 1234  
 QY 1041 ttaattggaaccttgatataaatttttagctgtgtacacgttcggcggtcctgtaattc 1100  
 |||||  
 Db 1235 ttaattggaaccttgatataaatttttagctgtgtacacgttcggcggtcctgtaattc 1294  
 QY 1101 agcctgtcacccgaatgctgactgtatgaatgagacagctcactacatgacccactctcct 1160  
 |||||  
 Db 1295 agcctgtcacccgaatgctgactgtatgaatgagacagctcactacatgacccactctcct 1354  
 QY 1161 cccactgaaggtcttcacgggctccaggtgagaccaaaggatgcacagcgctgcgcatg 1220  
 |||||  
 Db 1355 cccactgaaggtcttcacgggctccaggtgagaccaaaggatgcacagcgctgcgcatg 1414  
 QY 1221 ccccgagggcagcgttaagagatccaagatctagatttggttttagtcatgaataataa 1280  
 |||||  
 Db 1415 ccccgagggcagcgttaagagatccaagatctagatttggttttagtcatgaataataa 1474  
 QY 1281 acagcttcaaacactgcacaatttttcccccctttgaaagccactggggccaattgtgtg 1340  
 |||||  
 Db 1475 acagcttcaaacactgcacaatttttcccccctttgaaagccactggggccaattgtgtg 1534  
 QY 1341 ttaagaggtggtgtagataaagaatgtgaacgttgacatcttgcaggttgcagaagaatcc 1400  
 |||||  
 Db 1535 ttaagaggtggtgtagataaagaatgtgaacgttgacatcttgcaggttgcagaagaatcc 1594  
 QY 1401 aagcaggtatgtgcttagtgttaagggctttagacagcgttaaatatgaggaagaagt 1460  
 |||||  
 Db 1595 aagcaggtatgtgcttagtgttaagggctttagacagcgttaaatatgaggaagaagt 1654  
 QY 1461 ggcacagttagacatctgcagagatacaatctgagaggtcttcgttttcgcacttcgcaca 1520  
 |||||  
 Db 1655 ggcacagttagacatctgcagagatacaatctgagaggtcttcgttttcgcacttcgcaca 1714  
 QY 1521 gagctaggtcccttgatcttctttagatgtgaagtcgtctcttgacaacataatttgt 1580  
 |||||  
 Db 1715 gagctaggtcccttgatcttctttagatgtgaagtcgtctcttgacaacataatttgt 1774

QY 1581 aaaagttagtagtcttctttaaatacatataaagaagcttcgta 1625  
 |||||  
 Db 1775 aaaagttagtagtcttctttaaatacatataaagaagcttcgta 1819  
 RESULT 4  
 ID 236837  
 XX 236837 standard; cDNA: 522 BP.  
 AC 236837;  
 XX  
 DT 13-MAR-2000 (first entry)  
 XX  
 DE Partial nucleotide sequence of human interleukin-22.  
 XX  
 KW Human; interleukin-22; IL-22; IL-21; immune system disorder;  
 KW immune cell chemotaxis; haematopoietic cell disorder;  
 KW haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;  
 KW respiratory problem; organ rejection; graft-versus-host disease; GVHD;  
 KW inflammation; hyperproliferative disorder; tissue regeneration;  
 KW embryonic stem cell differentiation; embryonic stem cell proliferation;  
 KW haematopoietic lineage; allergic asthma; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..522  
 FT /tag= a  
 FT /product= "partial IL-22"  
 PN MO9961617-A1.  
 XX  
 PD 02-DEC-1999.  
 XX  
 PF 27-MAY-1999; 99WO-US11644.  
 XX  
 PR 29-MAY-1998; 98US-0087340.  
 PR 10-SEP-1998; 98US-0099805.  
 PR 30-APR-1999; 99US-0131965.  
 XX  
 PA (HDMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Ruben SM, Ebner R;  
 XX  
 DR WPI: 2000-072622/06.  
 DR P-PSDB: Y53893.  
 XX  
 PT Novel polynucleotides used to develop products for treating e.g. immune  
 PT disorders, blood disorders, autoimmune disorders, allergies,  
 PT inflammation, hyperproliferative disorders or infections  
 XX  
 PS Disclosure; Fig 8; 170pp; English.  
 XX  
 CC The present sequence encodes a partial human interleukin-22 (IL-22)  
 CC protein. The specification also describes IL-21 polynucleotides and  
 CC polypeptides. The IL-21 polynucleotide was isolated from a cDNA library  
 CC of epileptic frontal cortex. IL-21 and IL-22 may be useful in treating  
 CC deficiencies or disorders of the immune system, by activating or  
 CC inhibiting the proliferation, differentiation, or mobilization  
 CC (chemotaxis) of immune cells, treating or detecting deficiencies  
 CC of haematopoietic cells, to modulate haemostatic or thrombolytic  
 CC activity, in treating or detecting autoimmune disorders, treating  
 CC asthma (particularly allergic asthma) or other respiratory problems,  
 CC to treat and/or prevent organ rejection or graft-versus-host disease  
 CC (GVHD), to modulate inflammation, to treat or detect hyperproliferative  
 CC disorders, to treat or detect infectious agents, to differentiate,  
 CC proliferate and attract cells, leading to the regeneration of tissues,  
 CC IL-21 and IL-22 may also increase or decrease the differentiation or  
 CC proliferation of embryonic stem cells and haematopoietic lineage, may  
 CC be used to modulate mammalian characteristics.  
 CC  
 CC Sequence 522 BP; 77 A; 208 C; 169 G; 68 T; 0 other;

	Query Match	28.2%	Score 463.4;	DB 21;	Length 522;	
	Best Local Similarity	99.8%	Fred. NO. 4.4e-93;			
	Matches 464;	Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;
OY	gcccgtcctcagtgcttcacaccacacgcgtcgacgcgtgggccccgtgtgaagaagcgccgaac	80				
Dd						
	58 ggcgtgcctcagtgcttcacaccacacgcgtcgacgcgtgggccccgtgtgaagaagcgccgaac	117				
OY	81 ggcagctgcgccgagcaaggaggcagcccgccgaacgcgcgccttcgggccccacaactg	140				
Dd						
	118 gcgaagtgcgccgaggggagcagcccgccgaacgcgcgccttcgggccccacaactg	177				
OY	141 cgcgaacgtgtgcaccttgggacctaaagaatctctctaagaccgcggaggttaaccacggtac	200				
Dd						
	178 cgcaagcgtgtgcaccttgggacctaaagaatctctctaagaccgcggaggttaaccacggtac	237				
OY	201 ctgcctgaagaacctactgtcgtgtgcggggctgctctgaaccgggctgttcggcgagagagac	260				
Dd						
	238 ctgcctgaagaacctactgtcgtgtgcggggctgctctgaaccgggctgttcggcgagagagac	297				
OY	261 gtgcgtctccgacgagccccctgtttacaatgtccaacgcgtgtctctgcgcgcgaacccccgcc	320				
Dd						
	298 gtgcgtctccgacgagccccctgtttacaatgtccaacgcgtgtctctgcgcgcgaacccccgcc	357				
OY	321 tgcgcgcggcgagcgttccgcttaccacccgaagacctactgtacatcccccgtgggtgtgaccc	380				
Dd						
	358 tgcgcgcggcgagcgttccgcttaccacccgaagacctactgtacatcccccgtgggtgtgaccc	417				
OY	381 tgcgtcccccagcgagagagagacgacgacgacatccaactccagatcgacacaacagggc	440				
Dd						
	418 tgcgtcccccagcgagagagagacgacgacgacatccaactccagatcgacacaacagggc	477				
OY	441 gccaaagcttcctgtctgtggcccccaagcagcgccgcgtgtggcccttga	485				
Dd						
	478 gccaaagcttcctgtctgtggcccccaagcagcgccgcgtgtggcccttga	522				
RESULT	5					
ID	Z52198 standard; cdDNA: 1221 BP.					
XX						
AC	Z52198;					
XX						
DT	18-JUL-2000 (first entry)					
XX						
DE	Murine transforming growth factor beta-9; Ztfibeta-9 CDNA.					
XX						
KW	Murine transforming growth factor beta-9; Ztfi beta-9;					
KW	Alzheimer's disease; neurodegenerative disease; Huntington's disease;					
KW	amyotrophic lateral sclerosis; ALS; Parkinson's disease;					
KW	peripheral neuropathy; demyelinating disease; multiple sclerosis;					
KW	antiviral; cyostatic; ss.					
XX						
DS	Mus sp.					
XX						
FH	Key	Location/Qualifiers				
FF	CDS	79..696				
FT		/tag= a				
FT		/product= "Murine Ztfibeta-9 protein"				
FT	sig_peptide	79..144				
FT		/tag= b				
FT	mat_peptide	145..693				
FT		/tag= C				
FT		/product= "Mature Ztfibeta-9 protein represented as				
XX		SEQ ID NO:12"				
PN	WO200015798-A2.					
XX						
PD	23-MAR-2000.					
XX						
PF	17-SEP-1999;	99WO-US21677.				
XX						

PR 17-SEP-1998; 9805-0154617.  
PA (Zymo) ZYMOGENETICS INC.  
XX  
XX Presnell SR, Taft DW, Foley KP;  
PI WPL: 2000-271436/23.  
DR P-PsDB: Y70657.  
XX  
XX Polynucleotides encoding a novel transforming growth factor beta-9  
PT polypeptide, designated Ztfg beta-9, useful as an antiviral and  
PT antiproliferative agent -  
XX  
PS Disclosure; Page 86-88; 97pp; English.

XX  
XX The present sequence encodes murine transforming growth factor beta-9  
CC designated Ztfg beta-9. Murine Ztfg beta-9 was found to be highly  
CC expressed in the RGL hypothalamic cell line. This can be used  
CC to treat a variety of neurodegenerative diseases such as amyotrophic  
CC lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,  
CC Parkinson's disease and peripheral neuropathies, or demyelinating  
CC diseases including multiple sclerosis. Ztfg beta-9 peptides have  
CC antiviral activity and may also be used to regulate the proliferation,  
CC differentiation and apoptosis of neurons, glial cells, lymphocytes,  
CC hematopoietic cells and stromal cells.  
XX

SQ Sequence 1221 BP; 224 A; 378 C; 353 G; 266 T; 0 other:

[illegible]







PF 27-MAY-1999; 99WO-US11644.  
 XX  
 PR 29-MAY-1998; 98US-0087340.  
 PR 10-SEP-1998; 98US-0099805.  
 PR 30-APR-1999; 99US-0131965.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PI Ruben SM, Ebner R;  
 XX WPI: 2000-072622/06.  
 DR  
 XX  
 PT Novel polynucleotides used to develop products for treating e.g. immune  
 PT disorders, blood disorders, autoimmune disorders, allergies,  
 PT inflammation, hyperproliferative disorders or infections  
 XX  
 PS Disclosure: Page 163-164; 170pp; English.  
 XX  
 CC 236852-55 represent cDNA clones that are related to cDNA encoding human  
 CC interleukin-21 (IL-21) protein. The specification also describes  
 CC IL-22 polynucleotides and polypeptides. The IL-21 polynucleotide was  
 CC isolated from a cDNA library of apoptotic T-cells. IL-21 and IL-22 may  
 CC be useful in treating deficiencies or disorders of the immune system, by  
 CC activating or inhibiting the proliferation, differentiation, or  
 CC mobilization (chemotaxis) of immune cells, treating or detecting  
 CC deficiencies or disorders of haematopoietic cells, to modulate  
 CC haemostatic or thrombolytic activity, in treating or detecting  
 CC autoimmune disorders, treating asthma (particularly allergic asthma)  
 CC or other respiratory problems, to treat and/or prevent organ rejection  
 CC or graft-versus-host disease (GVHD), to modulate inflammation, to  
 CC treat or detect hyperproliferative disorders, to treat or detect  
 CC infectious agents, to differentiate, proliferate and attract cells,  
 CC leading to the regeneration of tissues, IL-21 and IL-22 may also  
 CC increase or decrease the differentiation or proliferation of embryonic  
 CC stem cells and haematopoietic lineage, may be used to modulate mammalian  
 CC characteristics.  
 CC  
 SO Sequence 264 BP; 52 A; 90 C; 77 G; 41 T; 4 other;

Query Match: 10.1%; Score 165.4; DB 21; Length 264;  
 Best Local Similarity 87.3%; Pred. No. 1.1e-27;  
 Matches 227; Conservative 0; Mismatches 24; Indels 9; Gaps 4;

QY 439 gccgaagctctgtctg-gcccaagacgagccgctgcccctgaagccgg-tctcg 496  
 DB 5 gaggcaagctctgtctgagcccaagacgagccgctgcccctgaagccgg-tctcg 64  
 QY 497 ccccggaagctctcccgccgagccgagccgagccgagctgagccgctgagggc 556  
 DB 65 ccccggaagctctcccgccgagccgagccgagccgagctgagccgctgagggc 124  
 QY 557 tccgttcgagcactctgaagagatg-----caccgagcaaaccaagtcgagcagc 609  
 DB 125 ttcggtgncgagcactctgaagagagatgcccagcagcaacgaagtcgagcagc 184  
 QY 610 cagcgccgcttccatgagagctgtaagcagctcatcctacacagcgagcctgctgct 669  
 DB 185 cagcgccgcttccatgagagctgtaagcagctcatcctacacagcgagcagcagcagc 244  
 QY 670 tgccttaagctacaagcaag 689  
 DB 245 tgccttaagctacaagcaag 264

RESULT 10  
 X89401  
 ID X89401 standard; cDNA; 540 BP.  
 XX X89401;  
 AC  
 XX 25-OC7-1999 (first entry)  
 DT  
 XX

DE Human interleukin-17D polynucleotide.  
 XX  
 KM Interleukin-17D; IL-17D; isoelectric point; cytokine; B cell tumour; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 XX CDS 1..540  
 FT /tag= a  
 FT /product= "IL-17D"  
 FT /note= "Interleukin-17D"  
 FT sig\_peptide 1..60  
 FT /tag= b  
 FT mat\_peptide 61..537  
 FT /tag= c  
 XX  
 XX WO9935267-A1.  
 PN  
 XX  
 PD 15-JUL-1999.  
 XX  
 XX 08-JAN-1999; 99WO-US00513.  
 PF  
 XX 09-JAN-1998; 98US-0070886.  
 PR  
 XX (IMV) IMMUNEX CORP.  
 PA  
 XX  
 XX Spriggs M, Upton C;  
 PI  
 XX WPI: 1999-478835/40.  
 DR P-PSDB; Y28235.  
 XX  
 PT Murine and Human interleukin 17D DNA, polypeptides and its  
 PT fragments, useful as molecular weight markers  
 XX  
 PS Claim 1: Page 6-7; 72pp; English.  
 XX  
 CC This sequence is the human interleukin-17D (IL-17D) polynucleotide.  
 CC IL-17D proteins are useful for determining the molecular weight of a  
 CC sample protein. Human IL-17D polypeptides are significantly related to  
 CC human IL-17 polypeptides. The homology between IL-17 and IL-17D suggests  
 CC that the IL-17D polypeptide is capable of signalling through cytokine  
 CC receptors. The IL-17D protein and fragments of it are useful as controls  
 CC for peptide fragmentation which can be used to determine the isoelectric  
 CC point of a sample protein. Antibodies generated against IL-17D and its  
 CC fragmented peptides can be used to enhance the accuracy of these  
 CC molecular weight markers. IL-17D can also be used as a therapeutic agent  
 CC for the treatment of diseases mediated by IL-17D. IL-17D polypeptides  
 CC bind to B cells. It is likely that these polypeptides can be used for  
 CC targeting compounds to B cells and B cell tumours, and for specific  
 CC selection of B cell populations.  
 CC  
 SO Sequence 540 BP; 102 A; 174 C; 169 G; 95 T; 0 other;

Query Match: 4.5%; Score 73.2; DB 20; Length 540;  
 Best Local Similarity 58.1%; Pred. No. 2.7e-07;  
 Matches 151; Conservative 0; Mismatches 105; Indels 6; Gaps 1;

QY 133 ccaactcgagcagctgtgccttgagcctacagaaatctccacgagcgaggtacc 192  
 DB 278 ccaacagaagagagcctgtcccttgagcctacagaaatctccacgagcgaggtacc 337  
 QY 193 ccaagtaacctgctgaagcctactgctgagcgaggtgctgagcagcgaggtgctgagc 252  
 DB 338 ccgtgacctgagcgagcagcgaggtgctgagcgaggtgctgagcagcgaggtgctgagc 397  
 QY 253 aaggaagcgtgagcgttcgagcagcgagcgtgctgagcagcgaggtgctgagcagcgagc 306  
 DB 398 aaggaagcgtgagcgttcgagcagcgagcgtgctgagcagcgaggtgctgagcagcgagc 457  
 QY 307 gccgaccccgctgagcagcgagcgtgctgagcagcgagcagcgagcagcgagcagcgagc 366  
 DB 458 tctgcgcgacacgagcagcgagcagcgagcagcgagcagcgagcagcgagcagcgagcagc 517

```
OY      367 ccgtggcgtgcactgcgtc 386
        1 |||||
Db      518 ctgtggcgtgcactgcgtc 537

RESULT  11
234282
XX      234282 standard; cDNA; 687 BP.
AC      234282;
XX
XX      07-DEC-1999 (first entry)
DE      Human PRO1031 nucleotide sequence.
XX
XX      Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
KM      probe; blood coagulation disorder; cancer; cellular adhesion disorder;
KW      secreted protein; transmembrane protein; ss.
XX
XX      Homo sapiens.
PN      W09946281-A2.
PD      16-SEP-1999.
XX
PF      08-MAR-1999; 99WO-US05028.
XX
PR      10-MAR-1998; 98US-0077450.
PR      11-MAR-1998; 98US-0077632.
PR      11-MAR-1998; 98US-0077641.
PR      12-MAR-1998; 98US-0077645.
PR      13-MAR-1998; 98US-0077791.
PR      17-MAR-1998; 98US-0078004.
PR      20-MAR-1998; 98US-0040220.
PR      20-MAR-1998; 98US-0078886.
PR      20-MAR-1998; 98US-0078910.
PR      20-MAR-1998; 98US-0078936.
PR      25-MAR-1998; 98US-0078939.
PR      26-MAR-1998; 98US-0079294.
PR      27-MAR-1998; 98US-0079656.
PR      27-MAR-1998; 98US-0079663.
PR      27-MAR-1998; 98US-0079664.
PR      27-MAR-1998; 98US-0079689.
PR      27-MAR-1998; 98US-0079728.
PR      30-MAR-1998; 98US-0079786.
PR      30-MAR-1998; 98US-0079920.
PR      31-MAR-1998; 98US-0079923.
PR      31-MAR-1998; 98US-0080105.
PR      31-MAR-1998; 98US-0080107.
PR      31-MAR-1998; 98US-0080165.
PR      01-APR-1998; 98US-0080194.
PR      01-APR-1998; 98US-0080327.
PR      01-APR-1998; 98US-0080328.
PR      01-APR-1998; 98US-0080333.
PR      01-APR-1998; 98US-0080334.
PR      08-APR-1998; 98US-0080349.
PR      08-APR-1998; 98US-0081049.
PR      08-APR-1998; 98US-0081070.
PR      09-APR-1998; 98US-0081071.
PR      09-APR-1998; 98US-0081195.
PR      09-APR-1998; 98US-0081203.
PR      15-APR-1998; 98US-0081229.
PR      15-APR-1998; 98US-0081817.
PR      15-APR-1998; 98US-0081838.
PR      15-APR-1998; 98US-0081952.
PR      15-APR-1998; 98US-0081955.
PR      21-APR-1998; 98US-0082568.
PR      21-APR-1998; 98US-0082569.
PR      22-APR-1998; 98US-0082700.
PR      22-APR-1998; 98US-0082704.
PR      23-APR-1998; 98US-0082804.
PR      23-APR-1998; 98US-0082767.
PR      23-APR-1998; 98US-0082796.
```

```
PR      27-APR-1998; 98US-0083336.
PR      28-APR-1998; 98US-0083322.
PR      29-APR-1998; 98US-0083392.
PR      29-APR-1998; 98US-0083495.
PR      29-APR-1998; 98US-0083496.
PR      29-APR-1998; 98US-0083496.
PR      29-APR-1998; 98US-0083500.
PR      29-APR-1998; 98US-0083500.
PR      29-APR-1998; 98US-0083545.
PR      29-APR-1998; 98US-0083554.
PR      29-APR-1998; 98US-0083558.
PR      29-APR-1998; 98US-0083559.
PR      30-APR-1998; 98US-0083742.
PR      05-MAY-1998; 98US-0084366.
PR      06-MAY-1998; 98US-0084414.
PR      06-MAY-1998; 98US-0084414.
PR      07-MAY-1998; 98US-0084459.
PR      07-MAY-1998; 98US-0084500.
PR      07-MAY-1998; 98US-0084527.
PR      07-MAY-1998; 98US-0084537.
PR      07-MAY-1998; 98US-0084639.
PR      07-MAY-1998; 98US-0084640.
PR      07-MAY-1998; 98US-0084643.
PR      13-MAY-1998; 98US-0085323.
PR      13-MAY-1998; 98US-0085338.
PR      13-MAY-1998; 98US-0085339.
PR      15-MAY-1998; 98US-0085573.
PR      15-MAY-1998; 98US-0085579.
PR      15-MAY-1998; 98US-0085580.
PR      15-MAY-1998; 98US-0085582.
PR      15-MAY-1998; 98US-0085889.
PR      15-MAY-1998; 98US-0085977.
PR      15-MAY-1998; 98US-0085700.
PR      15-MAY-1998; 98US-0085700.
PR      18-MAY-1998; 98US-0085703.
PR      18-MAY-1998; 98US-0086023.
PR      22-MAY-1998; 98US-0086392.
PR      22-MAY-1998; 98US-0086414.
PR      22-MAY-1998; 98US-0086430.
PR      22-MAY-1998; 98US-0086486.
PR      28-MAY-1998; 98US-0087098.
PR      28-MAY-1998; 98US-0087106.
PR      28-MAY-1998; 98US-0087208.
PR      30-JUL-1998; 98US-0094651.
PR      11-SEP-1998; 98US-0100038.
XX
XX      (GETH ) GENENTECH INC.
XX
XX      Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
XX      WPI; 1999-551358/46.
XX      DR      P-PSDB; Y41762.
XX
XX      PT      New secreted and transmembrane polypeptides and their polynucleotides,
XX      PT      useful for treating blood coagulation disorders, cancers and cellular
XX      PT      adhesion disorders -
XX
XX      PS      Claim 2; Fig 197; 530pp; English.
XX
XX      The present invention describes secreted and transmembrane polypeptides
XX      CC      and their polynucleotides. The nucleotide sequences are useful as
XX      CC      sources of probes, primers, for chromosome mapping, and for generation
XX      CC      of antisense sequences. They can also be used to create transgenic
XX      CC      animals. The proteins can be used to treat a variety of diseases and
XX      CC      disorders, depending on their function. Diseases that may be treated
XX      CC      include blood coagulation disorders, cancers and cellular adhesion
XX      CC      disorders. They may also be used to raise antibodies. Z33691 to
XX      CC      Z34338, and Y41685 to Y41774 represent polynucleotide and polypeptide
XX      CC      sequence given in the exemplification of the present invention.
XX
XX      SQ      Sequence 687 BP; 142 A; 216 C; 209 G; 120 T; 0 other;
```

Query Match 4.5%; Score 73.2; DB 20; Length 687;  
Best Local Similarity 58.1%; Pred. No. 3e-07;











GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2001, 12:49:48 ; Search time 26.41 Seconds  
(without alignments)  
207.157 Million cell updates/sec

Title: US-09-320-713-4

Perfect score: 864  
Sequence: 1 NSARARAVLAPFHHTQLGP.....SIDKQGAKLLGPNDAAPG 160

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq\_36:\*

- 1: /SIDSI/gcgdata/geneseq/geneseq/AA1980.DAT:\*
- 2: /SIDSI/gcgdata/geneseq/geneseq/AA1981.DAT:\*
- 3: /SIDSI/gcgdata/geneseq/geneseq/AA1982.DAT:\*
- 4: /SIDSI/gcgdata/geneseq/geneseq/AA1983.DAT:\*
- 5: /SIDSI/gcgdata/geneseq/geneseq/AA1984.DAT:\*
- 6: /SIDSI/gcgdata/geneseq/geneseq/AA1985.DAT:\*
- 7: /SIDSI/gcgdata/geneseq/geneseq/AA1986.DAT:\*
- 8: /SIDSI/gcgdata/geneseq/geneseq/AA1987.DAT:\*
- 9: /SIDSI/gcgdata/geneseq/geneseq/AA1988.DAT:\*
- 10: /SIDSI/gcgdata/geneseq/geneseq/AA1989.DAT:\*
- 11: /SIDSI/gcgdata/geneseq/geneseq/AA1990.DAT:\*
- 12: /SIDSI/gcgdata/geneseq/geneseq/AA1991.DAT:\*
- 13: /SIDSI/gcgdata/geneseq/geneseq/AA1992.DAT:\*
- 14: /SIDSI/gcgdata/geneseq/geneseq/AA1993.DAT:\*
- 15: /SIDSI/gcgdata/geneseq/geneseq/AA1994.DAT:\*
- 16: /SIDSI/gcgdata/geneseq/geneseq/AA1995.DAT:\*
- 17: /SIDSI/gcgdata/geneseq/geneseq/AA1996.DAT:\*
- 18: /SIDSI/gcgdata/geneseq/geneseq/AA1997.DAT:\*
- 19: /SIDSI/gcgdata/geneseq/geneseq/AA1998.DAT:\*
- 20: /SIDSI/gcgdata/geneseq/geneseq/AA1999.DAT:\*
- 21: /SIDSI/gcgdata/geneseq/geneseq/AA2000.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	864	100.0	160	21	Partial amino acid
2	835	96.6	173	21	Partial amino acid
3	835	96.6	185	21	Mature human trans
4	835	96.6	186	21	Mature human trans
5	835	96.6	187	21	Mature human trans
6	835	96.6	187	21	Mature human trans
7	835	96.6	202	21	Mature human trans
8	835	96.6	209	21	Human transforming
9	707	81.8	183	21	Human transforming
10	707	81.8	205	21	Mature murine tran
11	303	35.1	57	21	Human transforming
12	298	34.5	54	21	Human transforming

13	197.5	22.9	425	21	Y44462	Human Interleukin
14	192	22.2	197	21	Y92238	Human Interleukin
15	192	22.2	197	21	Y44460	Human Interleukin
16	192	22.2	197	21	Y53892	Amino acid sequenc
17	192	22.2	206	21	Y44485	Human Interleukin
18	180	20.8	34	21	Y70661	Human transforming
19	177	20.5	158	20	W85646	Murine Zcyto7 matu
20	177	20.5	160	20	W85645	Murine Zcyto7 matu
21	177	20.5	180	20	Y28236	Mouse Interleukin
22	177	20.5	180	20	Y22198	Murine EDIRF I pro
23	177	20.5	180	20	W85617	Murine Zcyto7, Mu
24	176.5	20.4	408	21	Y44461	Human Interleukin
25	174	20.1	128	20	W85643	Human Zcyto7 matur
26	174	20.1	151	20	W85641	Human Zcyto7 matur
27	174	20.1	153	20	W85642	Human Zcyto7 matur
28	174	20.1	154	20	W85640	Human Zcyto7 matur
29	174	20.1	157	20	W85644	Human Zcyto7 matur
30	174	20.1	158	20	W85639	Human Zcyto7 matur
31	174	20.1	160	20	W85618	Human Zcyto7 matur
32	174	20.1	160	20	W85619	Human Zcyto7 matur
33	174	20.1	160	20	W85620	Human Zcyto7 matur
34	174	20.1	160	20	W85621	Human Zcyto7 matur
35	174	20.1	160	20	W85622	Human Zcyto7 matur
36	174	20.1	160	20	W85626	Human Zcyto7 matur
37	174	20.1	160	20	W85627	Human Zcyto7 matur
38	174	20.1	160	20	W85628	Human Zcyto7 matur
39	174	20.1	160	20	W85629	Human Zcyto7 matur
40	174	20.1	175	21	Y44484	Human Interleukin
41	174	20.1	180	20	Y41762	Human PRO1031 prot
42	174	20.1	180	20	Y28235	Human Interleukin
43	174	20.1	180	20	Y22197	Human EDIRF I prot
44	174	20.1	180	20	W97350	Interleukin-20, H
45	174	20.1	180	20	W85615	Human Zcyto7, Hom

#### ALIGNMENTS

RESULT 1	
ID	Y53891 standard; Protein: 160 AA.
XX	Y53891;
AC	13-MAR-2000 (first entry)
XX	
DT	
DE	Partial amino acid sequence of human Interleukin-22.
XX	
KW	Human: Interleukin-22; IL-22; IL-21; Immune system disorder;
KW	Immune cell chemotaxis; haematopoietic cell disorder;
KW	haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;
KW	respiratory problem; organ rejection; graft-versus-host disease; GVHD;
KW	inflammation; hyperproliferative disorder; tissue regeneration;
KW	embryonic stem cell differentiation; embryonic stem cell proliferation;
KW	haematopoietic lineage; allergic asthma.
XX	
OS	Homo sapiens.
XX	
FH	
FT	Key
FT	Modified-site
FT	Location/Qualifiers
FT	26
FT	/note= "potential N-linked glycosylation site"
FT	57..64
FT	/note= "conserved domain I"
FT	72..77
FT	/note= "conserved domain II"
FT	99..105
FT	/note= "conserved domain III"
FT	121..128
FT	/note= "conserved domain IV"
FT	139
FT	/note= "potential N-linked glycosylation site"
XX	
PN	W0961617-A1.

[illegible]

KW		haemostatic activity; thrombolytic activity; autoimmune disorder; asthma;
KM		respiratory problem; organ rejection; graft-versus-host disease; GVHD;
KM		inflammation; hyperproliferative disorder; tissue regeneration;
KV		embryonic stem cell differentiation; embryonic stem cell proliferation;
KW		haematopoietic lineage; allergic asthma.
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Domain	18..23 "conserved domain VI"
FT	Modified-site	/note- "39..41
FT	Domain	/note- "Asn39 is a potential N-linked glycosylation site"
FT	Domain	60..65 /note- "conserved domain VII"
FT	Domain	69..77 /note- "conserved domain I"
FT	Domain	85..90 /note- "conserved domain II"
FT	Domain	112..118 /note- "conserved domain III"
FT	Domain	134..141 /note- "conserved domain IV"
FT	Modified-site	152..154 /note- "Asn152 is a potential N-linked glycosylation site"
FT		
FN		
PN	WO9961617-A1.	
XX		
PD	02-DEC-1999.	
XX		
PE	27-MAY-1999;	99NO-US11644.
XX		
PR	29-MAY-1998;	98US-0087340.
PR	10-SEP-1998;	98US-0099805.
PR	30-APR-1999;	99US-0131965.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Ruben SM, Ebner R;	
DR	WPI: 2000-072622/06.	
DR	N-PsDB: 236837.	
XX		
PT	Novel polynucleotides used to develop products for treating e.g. immune disorders, blood disorders, autoimmune disorders, allergies,	
PT	inflammation, hyperproliferative disorders or infections -	
XX		
PS	Disclosure; Fig 8; 170pp; English.	
XX		
CC	The present sequence represents a partial human interleukin-22 (IL-22)	
CC	protein. The specification also describes IL-21 polynucleotides and	
CC	polypeptides. The IL-21 polynucleotide was isolated from a cDNA library	
CC	of epileptic frontal cortex. IL-21 and IL-22 may be useful in treating	
CC	deficiencies or disorders of the immune system, by activating or	
CC	inhibiting the proliferation, differentiation, or mobilization	
CC	(Chemotaxis) of immune cells, treating or detecting deficiencies	
CC	of haematopoietic cells, to modulate haemostatic or thrombolytic	
CC	activity, in treating or detecting autoimmune disorders, treating	
CC	asthma (particularly allergic asthma) or other respiratory problems,	
CC	to treat and/or prevent organ rejection or graft-versus-host disease	
CC	(GVHD), to modulate inflammation, to treat or detect hyperproliferative	
CC	disorders, to treat or detect infectious agents, to differentiate,	
CC	proliferate and attract cells, leading to the regeneration of tissues,	
CC	IL-21 and IL-22 may also increase or decrease the differentiation or	
CC	proliferation of embryonic stem cells and haematopoietic lineage, may	
CC	be used to modulate mammalian characteristics.	
XX		
SQ	Sequence 173 AA;	
Query Match	96.6%; Score 835; DB 21; Length 173;	
Best Local Similarity	98.7%; Ptd. No. 8, 2e-86;	

Query Match	96.68;	Score 835;	DB 21;	Length 185;
Best Local Similarity	98.78;	Pred. No. 8.9e-66;		
Matches 154;	Conservative	0;	Mismatches 2;	Indels 0;
			Gaps	0

Query Match	96.6%;	Score 835;	DB 21;	Length 186;
Best Local Similarity	98.7%;	Pred. No. 9e-86;		
Matches 154; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0



Db	47	aagvtsafhtlqigp	regarnscpaggrpadrrtrrptlnlrsvspwayrrisydcparyp	106
Oy	65	RYLPAYCLRCGCLTGL	FEDEDVRFRSAPVYMPTVVLRRTPACAGRSVYTEAVYTPWG	124
Db	107	rylpayclrcrgclglg	lfgedvdrfrsapytmptvllrrtpacagrsvlyteavtlpvg	166
Oy	125	CTCVPPEPEKDADSI	NSIDIKGAKLLIGENDAPACR	160
Db	167	ctcvpepekadsins	idsidkgakillspndapap	202
RESULT 8				
Y70662				
Y70662			standard; protein; 209 AA.	
Y70662:				
18-JUL-2000			(first entry)	
Human transforming growth factor beta-9, Ztgbeta-9 variant protein.				
Human transforming growth factor beta-9; Ztgbeta-9;				
Alzheimer's disease; neurodegenerative disease; Huntington's disease;				
amyotrophic lateral sclerosis; ALS; Parkinson's disease;				
peripheral neuropathy; demyelinating disease; multiple sclerosis;				
antiviral; cyostatic.				
Homo sapiens.				
Key		Location/Qualifiers		
Peptide		1..22		
		/label= signal_peptide		
WO200015798-A2.				
23-MAR-2000.				
17-SEP-1999;		99WO-US21677.		
17-SEP-1998;		98US-0154817.		
(ZYMO ) ZYMOGENETICS INC.				
Presnell SR, Taft DM, Foley KP;				
WP1: 2000-271436/23.				
N-PSDB; Z52201.				
Polynucleotides encoding a novel transforming growth factor beta-9				
polypeptide, designated Ztgbeta-9, useful as an antiviral and				
antiproliferative agent				
Claim 6; Page 92-93; 97pp; English.				
The present sequence is a variant of human transforming growth				
factor beta-9, designated Ztgbeta-9. Human Ztgbeta-9 was isolated				
from an arrayed pituitary gland cDNA plasmid library by PCR screening.				
This can be used to treat a variety of neurodegenerative diseases such				
as amyotrophic lateral sclerosis (ALS), Alzheimer's disease, Huntington's				
disease, Parkinson's disease and peripheral neuropathies, or				
demyelinating diseases including multiple sclerosis. Ztgbeta-9 peptides				
have antiviral activity and may also be used to regulate the				
proliferation, differentiation and apoptosis of neurons, glial cells,				
lymphocytes, hematopoietic cells and stromal cells.				
Sequence 209 AA;				
Query Match	96.6%;	Score 835;	DB 21;	Length 209;
Best Local Similarity	98.7%;	Pred. NO. 1e-85;		
Matches 154;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0

QY	5	ARAVLSAHHHTLQGLPREQARNA	SC	PAGGRPADRRFPPTNLR	RSVSPMAVRRISYDPARP	64
		1				
Db	54	aayvlsahhhtlqglpreqar	na	scpaggrpadrrfpptnl	rsvspmayrrisydparyp	113
QY	65	RYLPEAYLCRQCLTGLFGEEDV	RFRSA	PVYMPVTLVLRTPACAG	RSRYTEAVYTIPIVG	124
Db	114	rylpeaylcrqcltglfgeedv	rfrsapympvtlvlrtp	acagrsryteavtytipvg		173
QY	125	CTCYPEPEKDA	STNS	STDKGAKLLLPNDP	PAPG	160
Db	174	ctcypepekadas	ins	sldkqgaklllpndapagp	209	
RESULT	9					
ID	Y70658					
XX	Y70658	standard; Protein; 183	AA.			
XX	XX	Y70658;				
XX	XX	18-JUL-2000	(first entry)			
DE	XX	Mature murine transforming growth factor beta-9, Ztgfbeta-9 protein.				
XX	XX	Murine transforming growth factor beta-9; Ztgf beta-9;				
KW	XX	Alzheimer's disease; neurodegenerative disease; Huntington's disease;				
KM	XX	amyotrophic lateral sclerosis; ALS; Parkinson's disease;				
KW	XX	peripheral neuropathy; demyelinating disease; multiple sclerosis;				
XX	XX	antiviral; cytosolic.				
OS	XX	Mus sp.				
XX	XX	WO200015798-A2.				
PN	XX	23-MAR-2000.				
PD	XX	17-SEP-1999;	99WO-US21677.			
PF	XX	17-SEP-1998;	98US-0154817.			
PR	XX	(ZYMO ) ZYMOGENETICS INC.				
PA	XX	Presnell SR, Taft DW, Foley KP;				
PI	XX	WPI: 2000-271436/23.				
XX	XX	N-PSDB: 252198.				
DR	XX	Polynucleotides encoding a novel transforming growth factor beta-9				
PT	XX	polypeptide, designated ztgf beta-9, useful as an antiviral and				
PT	XX	antiproliferative agent				
PS	XX	Claim 6; Page 89; 97p; English.				
XX	XX	The present sequence is the mature murine transforming growth factor				
CC	XX	beta-9, designated ztgf beta-9. The signal sequence extends from				
CC	XX	amino acid residues 1 to 22. Murine ztgf beta-9 was found to be highly				
CC	XX	expressed in the HCL hypotlamic cell line. This can be used				
CC	XX	to treat a variety of neurodegenerative diseases such as amyotrophic				
CC	XX	lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,				
CC	XX	Parkinson's disease and peripheral neuropathies, or demyelinating				
CC	XX	diseases including multiple sclerosis. Ztgf beta-9 peptides have				
CC	XX	antiviral activity and may also be used to regulate the proliferation,				
CC	XX	differentiation and apoptosis of neurons, glial cells, lymphocytes,				
CC	XX	hematopoietic cells and stromal cells.				
XX	XX	Sequence	183	AA;		
XX	XX	Sequence	183	AA;		
QY	5	ARAVLSAHHHTLQGLPREQARNA	SC	PAGGRPADRRFPPTNLR	RSVSPMAVRRISYDPARP	64
		1				
		Query Match	81.8%;	Score 707;	DB 21;	Length 183;
		Best Local Similarity	84.5%;	Pred. No. 1.8e-71;		
		Matches 131;	Conservative 6;	Mismatches 14;	Indels 4;	Gaps 1

Db	QY	125	CTCYPEPEEKADSDINSSIDKQAKLLGPNAPAG	159	152	ctcvepekdadsanssmd----	Killipdrpag	182
Db	QY	32	aagylasfhtltqlgpreqarnascpagrreadrtrffipnllrsvspwayrlsydparfp	91				
QY	65	RYLPEAVCLRGCLTGLGFEEDVNRFSAPVYMPVVLRRPPACGSGSVTEAVTTPVG	124					
Db	92	rylpeayclrgcltglygeedfrftrspvlpavllrrtacaggrsvyaehtltpvg	151					
Db	152	ctcvepekdadsanssmd----	Killipdrpag	182				
RESULT	10							
ID	Y70657	Y70657	standard; Protein; 205 AA.					
AC	Y70657;							
DT	18-JUL-2000	(first entry)						
DE	Murine transforming growth factor beta-9, Ztgf beta-9 protein.							
KW	Murine transforming growth factor beta-9; Ztgf beta-9;							
KW	Alzheimer's disease; neurodegenerative disease; Huntington's disease;							
KW	amyotrophic lateral sclerosis; ALS; Parkinson's disease;							
KW	peripheral neuropathy; demyelinating disease; multiple sclerosis;							
KX	antiviral; cytosolic.							
OS	Mus sp.							
FT	Key	Location/Qualifiers						
FT	Peptide	1..22						
XX		/label= Signal_peptide						
XX	MO200015798-A2.							
PD	23-MAR-2000.							
XX	17-SEP-1999;	99WO-US21677.						
XX	17-SEP-1998;	98US-0154817.						
XX	(Zymo ) ZYMOGENETICS INC.							
XX	Presnell SR, Taft DW, Foley KP;							
XX	WPI: 2000-271436/23.							
XX	N-PSDB: 252198.							
XX	Polynucleotides encoding a novel transforming growth factor beta-9							
XX	polypeptide, designated Ztgf beta-9, useful as an antiviral and							
XX	antiproliferative agent -							
XX	Claim 6; Page 88; 97pp: English.							
XX	The present sequence encodes murine transforming growth factor beta-9,							
XX	designated Ztgf beta-9. Murine Ztgf beta-9 was found to be highly							
XX	expressed in the HCL hypochalaemic cell line. This can be used							
XX	to treat a variety of neurodegenerative diseases such as amyotrophic							
XX	lateral sclerosis (ALS), Alzheimer's disease, Huntington's disease,							
XX	Parkinson's disease and peripheral neuropathies, or demyelinating							
XX	diseases including multiple sclerosis. Ztgf beta-9 peptides have							
XX	antiviral activity and may also be used to regulate the proliferation,							
XX	differentiation and apoptosis of neurons, glial cells, lymphocytes,							
XX	hematopoietic cells and stromal cells.							
XX	Sequence	205 AA;						
XX	Query Match	81.8%; Score 707; DB 21; Length 205;						
XX	Best Local Similarity	84.5%; Pred. No. 2,1e-71;						
XX	Matches 131; Conservative	6; Mismatches 14; Indels 4; Gaps						
XX	5 ARAVLSAFHHTLQLGPREQARNASC	PAGGRADRRFRPPTNLRSVSPWAYRLSYDPARFP	64					

Db 54 aagvisafhtclqgpreqarnascpagradrirfipnlrsvspwayrslsydparfp 113  
 QY 65 RFLPEAYCCRCGLTGLNEEDVFRFSAPVYMWPTVLRTPACAGRSYTYEAYTIPG 124  
 Db 114 rylpeaycicrgclqglgedftrfsvpavlrrtaacagrsyaehtlipyg 173  
 QY 125 CTCVPEPEKADSIINSIDKOGAKLLGPNDAAG 159  
 Db 174 ctcvpepkadsansmd----kllgpadrpag 204

RESULT 11  
 ID Y70666 standard; Protein: 57 AA.  
 AC Y70666;

18-JUL-2000 (first entry)

DE Human transforming growth factor beta-9, ztgfbeta-9 peptide-6.

KW Human transforming growth factor beta-9; Ztgf beta-9;  
 KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
 KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
 KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
 KW antiviral; cyostatic.

OS Homo sapiens.  
 PN WO200015798-A2.

23-MAR-2000.

PF 17-SEP-1999; 99WO-US21677.

PR 17-SEP-1998; 98US-0154817.

PA (ZYMO) ZYMOGENETICS INC.

PI Presnell SR, Taft DW, Foley KP;

DR WPI: 2000-271436/23.

PT Polynucleotides encoding a novel transforming growth factor beta-9

PT polypeptide, designated Ztgf beta-9, useful as an antiviral and

PS antiproliferative agent

CC Claim 6; Page 94-95; 97pp; English.

CC The present sequence is an epitope-bearing peptide derived from the human  
 CC transforming growth factor beta-9, designated Ztgf beta-9. These peptides  
 CC are used to raise antibodies, including anti-idiotypic antibodies. This  
 CC can be used to treat a variety of neurodegenerative diseases such as  
 CC amyotrophic lateral sclerosis (ALS), Alzheimer's disease, Huntington's  
 CC disease, Parkinson's disease and peripheral neuropathies, or  
 CC demyelinating diseases including multiple sclerosis. Ztgf beta-9 peptides  
 CC have antiviral activity and may also be used to regulate the  
 CC proliferation, differentiation and apoptosis of neurons, glial cells,  
 CC lymphocytes, hematopoietic cells and stromal cells.

CC Sequence 57 AA;

Query Match 35.1%; Score 303; DB 21; Length 57;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-27;  
 Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 102 RRTPCAGRSYTYEAYTIPVGCCTVPEPEKADSIINSIDKOGAKLLGPNDAAG 158  
 Db 1 rrtpcagrsytyeaytllpvgctcvpepkadsinsidkgakllgpnadapa 57

RESULT 12  
 ID Y70664 standard; Protein: 54 AA.  
 AC Y70664;

18-JUL-2000 (first entry)

DE Human transforming growth factor beta-9, ztgfbeta-9 peptide-4.

KW Human transforming growth factor beta-9; Ztgf beta-9;  
 KW Alzheimer's disease; neurodegenerative disease; Huntington's disease;  
 KW amyotrophic lateral sclerosis; ALS; Parkinson's disease;  
 KW peripheral neuropathy; demyelinating disease; multiple sclerosis;  
 KW antiviral; cyostatic.

OS Homo sapiens.  
 PN WO200015798-A2.

23-MAR-2000.

PF 17-SEP-1999; 99WO-US21677.

PR 17-SEP-1998; 98US-0154817.

PA (ZYMO) ZYMOGENETICS INC.

PI Presnell SR, Taft DW, Foley KP;

DR WPI: 2000-271436/23.

PT Polynucleotides encoding a novel transforming growth factor beta-9

PT polypeptide, designated Ztgf beta-9, useful as an antiviral and

PS antiproliferative agent

CC Claim 6; Page 94; 97pp; English.

CC The present sequence is an epitope-bearing peptide derived from the human  
 CC transforming growth factor beta-9, designated Ztgf beta-9. These peptides  
 CC are used to raise antibodies, including anti-idiotypic antibodies. This  
 CC can be used to treat a variety of neurodegenerative diseases such as  
 CC amyotrophic lateral sclerosis (ALS), Alzheimer's disease, Huntington's  
 CC disease, Parkinson's disease and peripheral neuropathies, or  
 CC demyelinating diseases including multiple sclerosis. Ztgf beta-9 peptides  
 CC have antiviral activity and may also be used to regulate the  
 CC proliferation, differentiation and apoptosis of neurons, glial cells,  
 CC lymphocytes, hematopoietic cells and stromal cells.

CC Sequence 54 AA;

Query Match 34.5%; Score 298; DB 21; Length 54;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-26;  
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 QUGPREQARNASCPCGGRPADRRFRPPNLRKSVPMATYISTDPARYPRIDEA 70  
 Db 1 qugpreqarnascpcggrpadrrfrrpnlrsvpmatyistdparypridep 54

RESULT 13  
 ID Y44462 standard; Protein: 425 AA.  
 AC Y44462;

27-MAR-2000 (first entry)

DE Human Interleukin 17C-IgG1 Fc fusion protein, hIL-17C.fc.

KW Interleukin; IL-17C.fc; fusion protein; PRO1122 polypeptide; cytokine;  
 KW human IgG1; fluorescence-activated cell sorter analysis; FACS;





Db 183 lhvpgctcv 192

RESULT 15

Y44460

ID Y44460 standard; Protein: 197 AA.

AC Y44460;

XX

XX

DT 27-MAR-2000 (first entry)

XX

DE Human interleukin 17C, PRO1122 polypeptide.

XX

XX Interleukin: IL-17C: PRO1122 polypeptide: clone DNA62377-1381-1; UNQ561.

KW cytokine IL-17; cytotoxic T-lymphocyte-associated antigen 8; CTLA-8;

KW hybridisation probe; antagonist; degenerative cartilaginous disorder;

KW agonist; diagnose; therapy.

XX

OS Homo sapiens.

XX

XX

XX Key Location/Qualifiers

FT Peptide 1..18

FT /label= Signal\_peptide

FT Protein 19..197

FT /label= Mature\_IL-17C\_polypeptide

FT /note= "used to treat degenerative cartilaginous disorder"

FT Misc-difference 109

FT /note= "Conserved Trp residue"

FT Misc-difference 129

FT /note= "Conserved Cys residue"

FT Misc-difference 134

FT /note= "Conserved Cys residue"

FT Misc-difference 163

FT /note= "Conserved Cys residue"

FT Misc-difference 189

FT /note= "Conserved Cys residue"

FT Misc-difference 191

FT /note= "Conserved Cys residue"

XX

XX WO9960127-A2.

PN

XX

XX

PD 25-NOV-1999.

XX

XX

PF 14-MAY-1999; 99WO-US10733.

XX

XX

PR 15-MAY-1998; 98US-0085579.

PR 23-DEC-1998; 98US-0113621.

XX

XX

PA (GENE ) GENENTECH INC.

XX

XX

PI Chen J, Filvaroff E, Goddard A, Gurney AL, Li H, Wood WI;

XX

DR WPI: 2000-116314/10.

DR N-PSDB; 229728.

XX

XX

PT New polypeptides designated PRO1031 and PRO1122 used to treat a

PT degenerative cartilaginous disorder

XX

XX

PS Claim 23; Fig 3; 141pp; English.

XX

XX

CC The present sequence is the human PRO1122 polypeptide, also referred to

CC as UNQ561, and as Interleukin-17C (IL-17C), encoded by

CC clone DNA62377-1381-1. This sequence has identity with the

CC cytokine IL-17 and cytotoxic T-lymphocyte-associated antigen 8 (CTLA-8)

CC and has leucine zipper pattern. PRO1122 is expressed in pancreas, small

CC intestine, stomach and testis also. It shares about 26-28% amino acid

CC identity with IL-17 and IL-17B. The entire coding region of IL-17C can

CC be used as hybridisation probe. The PRO1122 polypeptide, agonist or

CC antagonist, is used to diagnose and treat a degenerative cartilaginous

CC disorder.

XX

XX

Sequence 197 AA:

Query Match 22.2%; Score 192; DB 21; Length 197;

Best Local Similarity 34.6%; Pred. No. 8,4e-14;

Matches 45; Conservative 17; Mismatches 58; Indels 10; Gaps 4;

QY 7 AVLSAFHHHTLQIGPRBO-ARNASCPAGGRPADRRFRPPTNLRSVPMAYRISYDPAPRPR 65

DB 65 alvsaleaashrghrhpsttcgv-llp-eevleadtqrsispwryrvdtddrtpq 122

QY 66 YLPEAYCLNGCLTGLFGEEDVFRSAPVYMPVVLRRTPACAGRSVYTEA-----Y 118

DB 123 klafaeclorgcidartgtretaalnsvrlqlslvlrrrrpcsrdsqglpqpafafhtet 182

QY 119 VTIPVGCTCY 128

DB 183 lhvpgctcv 192

Search completed: February 16, 2001, 12:49:48

Job time: 113 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2001, 12:50:10 ; Search time 18.64 Seconds  
(without alignments)  
154.138 Million cell updates/sec

Title: US-09-320-713-4

Perfect score: 864

Sequence: 1 NSARARAVLSAFHHITQLGP.....SIDKQAKLLGPNDAPAGP 160

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 174772 seqs, 17957048 residues

Total number of hits satisfying chosen parameters: 174772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: \*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep: \*  
3: /cgn2\_6/ptodata/2/1aa/6.COMB.pep: \*  
4: /cgn2\_6/ptodata/2/1aa/PCTUS.COMB.pep: \*  
5: /cgn2\_6/ptodata/2/1aa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	159	18.4	153 1 US-08-514-014-12	Sequence 12, Appl
2	159	18.4	153 2 US-08-833-823-12	Sequence 12, Appl
3	159	18.4	163 3 US-09-034-810-2	Sequence 2, Appl
4	159	18.4	163 3 US-08-685-239-2	Sequence 2, Appl
5	112	13.0	158 2 US-08-620-694A-7	Sequence 7, Appl
6	112	13.0	158 3 US-09-022-255-7	Sequence 7, Appl
7	112	13.0	158 3 US-09-022-696-7	Sequence 7, Appl
8	112	13.0	158 3 US-09-022-253-7	Sequence 7, Appl
9	112	13.0	158 3 US-09-022-260-7	Sequence 7, Appl
10	108.5	12.6	151 2 US-08-620-694A-8	Sequence 8, Appl
11	108.5	12.6	151 3 US-09-034-810-6	Sequence 8, Appl
12	108.5	12.6	151 3 US-09-022-255-8	Sequence 8, Appl
13	108.5	12.6	151 3 US-09-022-696-8	Sequence 8, Appl
14	108.5	12.6	151 3 US-08-685-239-6	Sequence 6, Appl
15	108.5	12.6	151 3 US-09-022-253-8	Sequence 6, Appl
16	108.5	12.6	151 3 US-09-022-260-8	Sequence 8, Appl
17	108	12.5	150 3 US-09-034-810-4	Sequence 8, Appl
18	108	12.5	150 3 US-08-685-239-4	Sequence 4, Appl
19	77.5	9.0	164 1 US-08-154-915-6	Sequence 4, Appl
20	75.5	8.7	397 4 PCT-US94-09700-11	Sequence 11, Appl
21	75	8.7	361 2 US-08-483-926A-9	Sequence 9, Appl
22	71.5	8.3	164 1 US-07-970-462A-2	Sequence 2, Appl
23	71.5	8.3	164 1 US-08-406-248-2	Sequence 2, Appl
24	71.5	8.3	164 1 US-08-574-03A-2	Sequence 2, Appl
25	71.5	8.3	164 2 US-08-524-218A-2	Sequence 2, Appl
26	71.5	8.3	164 2 US-08-795-015-2	Sequence 2, Appl
27	71.5	8.3	164 4 PCT-US92-10904-2	Sequence 2, Appl
28	71.5	8.3	164 4 PCT-US94-09700-2	Sequence 2, Appl

29	71.5	8.3	164 4 PCT-US94-12936-2	Sequence 2, Appl
30	71.5	8.3	164 4 PCT-US96-11886A-2	Sequence 2, Appl
31	70.5	8.2	120 5 5514582-36	Patent No. 5514582
32	70.5	8.2	164 1 US-08-275-983B-4	Sequence 4, Appl
33	70	8.1	299 3 US-09-188-930-192	Sequence 192, App
34	70	8.1	299 3 US-09-188-930-332	Sequence 332, App
35	69	8.0	362 2 US-08-737-045-13	Sequence 13, Appl
36	68	7.9	2205 1 US-08-093-453B-2	Sequence 2, Appl
37	67	7.8	199 3 US-09-188-930-250	Sequence 290, App
38	66	7.6	676 1 US-08-343-785-2	Sequence 2, Appl
39	66	7.6	676 1 US-08-343-785-8	Sequence 8, Appl
40	66	7.6	676 2 US-08-462-221-2	Sequence 2, Appl
41	66	7.6	676 2 US-08-462-221-8	Sequence 8, Appl
42	66	7.6	676 3 US-08-946-458-2	Sequence 2, Appl
43	66	7.6	676 3 US-08-946-458-8	Sequence 8, Appl
44	65	7.5	149 1 US-08-469-427A-14	Sequence 14, Appl
45	65	7.5	149 2 US-08-039-297B-2	Sequence 2, Appl

#### ALIGNMENTS

```
RESULT 1
US-08-514-014-12
; Sequence 12, Application US/08514014
; Patent No. 5707829
;
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: Kelleher, Kerry
; APPLICANT: Carlin, McKenough
; TITLE OF INVENTION: DNA SEQUENCES AND SECRETED PROTEINS
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESS: Genetics Institute, Inc. -- Legal Affairs
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02140
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/514,014
; FILING DATE:
;
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; REFERENCE/DOCKET NUMBER: G16000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 153 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-514-014-12
;
; Query Match 18.4%; Score 159; DB 1; Length 153;
; Best Local Similarity 32.6%; Pred. No. 6.4e-11;
; Matches 46; Conservative 18; Mismatches 53; Indels 24; Gaps 7;
;
; 1 NSARARAVLSAFHHITQLGP...PAGGRPAD-----RRFRPTNL--RS 48
; : | | : | | : | | : | | : | | : | | : | | : | | : | |
; Db 16 SNAARKIPKVTHTFTQ-----KPSGCPYGGSMKLDIGITINENQVSRVSRNIESRS 68
```



GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
APPLICANT: Kelleher, Kerry  
APPLICANT: Carlin, McKeough  
APPLICANT: Goldman, Samuel  
APPLICANT: Piltman, Debra  
APPLICANT: Mi, Sha  
APPLICANT: Neben, Steven  
APPLICANT: Giannotti, JoAnn  
APPLICANT: Golden/Fleet, Margaret  
TITLE OF INVENTION: Human CTLA-8 and Uses of CTLA-8-Related Proteins  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 Cambridgepark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/685,239  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Brown, Scott A.  
REGISTRATION NUMBER: 32,724  
REFERENCE/DOCKET NUMBER: GI5562  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8224  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 163 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-685-239-2

Query Match 18.4%; Score 159; DB 3; Length 163;  
Best Local Similarity 32.6%; Pred. No. 7e-11;  
Matches 46; Conservative 18; Mismatches 53; Indels 24; Gaps 7;

QY 1 NSARARAVLSAFHHTLQIGREQARNASC---PAGGRPAD-----RRRPPPTNL--RS 48  
DB 26 SEAAARKIPVGHFFD-----KPESCPVVGSKMKLDIGIINENORVMSRNIESRS 78  
QY 49 VSPFAYRISYDPARYPRYLPEAYCLRGCLTGLFGEDVAFRSAPYVMTVLRTTPACA 108  
DB 79 TSPMNTVTWMDPRPYPSEVYQAQCNLGCINAG-GKEDISMNSVPIQOELVVRKHQ-- 135  
QY 109 GGRSV-YTEAVYTIIVGCTCV 128  
DB 136 -GCSVSFOLEKVLVTGCTCV 155

RESULT 5  
US-08-620-694A-7  
Sequence 7, Application US/08620694A  
Patent No. 5869286  
GENERAL INFORMATION:  
APPLICANT: Yao, Zhengbin  
APPLICANT: Spriggs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 5869286e1 Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation

STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/620,694A  
FILING DATE: 21 MARCH 1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/538,765  
FILING DATE: 7 AUGUST 1995  
CLASSIFICATION: 435  
APPLICATION NUMBER: USSN 08/410,535  
FILING DATE: 23 MARCH 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 587-0430  
TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 158 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-620-694A-7

Query Match 13.0%; Score 112; DB 2; Length 158;  
Best Local Similarity 33.3%; Pred. No. 1.7e-05;  
Matches 33; Conservative 14; Mismatches 44; Indels 8; Gaps 5;

QY 32 GGRPADRRPPPTNL-RSVSPMAYRISYDPARYPRYLPEAYCLRGCLTGLFGEDVRF 90  
DB 58 GARVSSR--RPSDYLNKSSPWLHREDDPRYPYWEAQRHRCVNA-EGKLDHNM 114  
QY 91 SAPVYMTVLRRTP-ACAGRSVYTEAVYTIIVGCTCV 128  
DB 115 SVLIQDEILVLRKREPSCP---FTFRVKKLVGVGCTCV 150

RESULT 6  
US-09-022-255-7  
Sequence 7, Application US/09022255  
Patent No. 6072033  
GENERAL INFORMATION:  
APPLICANT: Yao, Zhengbin  
APPLICANT: Spriggs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 6072033e1 Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,255
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/620,694
FILING DATE: 21 MARCH 1996
APPLICATION NUMBER: USSN 08/538,765
FILING DATE: 7 AUGUST 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
FILING DATE: 23 MARCH 1995
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 158 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-022-255-7
```

```

Query Match          13.0%; Score 112; DB 3; Length 158;
Best Local Similarity 33.3%; Pred. No. 1.7e-05;
Matches 33; Conservative 14; Mismatches 44; Indels 8; Gaps 5;
```

```

Qy 32 GGRPADRRFRPTNL-RSVPMAVRIISYDPARYPRYLPEAYCLRCGLTGLGEEVDVFR 90
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 58 GAKVSSR--RPSDYLNRSSTPWLHNRNEDPDRYPSVIMWAOQRHRCVNA-EGKIDHHMN 114
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 91 SAPVYMPVTVLRRTP-ACAGGRSVYTEAVYTIIPVGCCTCV 128
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 115 SVLIQDELIVLKRPESCP---FTFRVEKMLVGVGCTCV 150
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```

RESULT 7
US-09-022-696-7
Sequence 7, Application US/09022696
Patent No. 6072037
GENERAL INFORMATION:
APPLICANT: Yao, Zhengbin
APPLICANT: Spriggs, Melanie
APPLICANT: Fanslow, William
TITLE OF INVENTION: No. 6072037e1 Receptor That Binds IL-17
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,696
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/620,694
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
```

```

FILING DATE: 23 MARCH 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
TELEFAX: (206)
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 158 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-022-696-7
```

```

Query Match          13.0%; Score 112; DB 3; Length 158;
Best Local Similarity 33.3%; Pred. No. 1.7e-05;
Matches 33; Conservative 14; Mismatches 44; Indels 8; Gaps 5;
```

```

Qy 32 GGRPADRRFRPTNL-RSVPMAVRIISYDPARYPRYLPEAYCLRCGLTGLGEEVDVFR 90
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 58 GAKVSSR--RPSDYLNRSSTPWLHNRNEDPDRYPSVIMWAOQRHRCVNA-EGKIDHHMN 114
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 91 SAPVYMPVTVLRRTP-ACAGGRSVYTEAVYTIIPVGCCTCV 128
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 115 SVLIQDELIVLKRPESCP---FTFRVEKMLVGVGCTCV 150
   | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```

RESULT 8
US-09-022-253-7
Sequence 7, Application US/09022253
Patent No. 6096305
GENERAL INFORMATION:
APPLICANT: Yao, Zhengbin
APPLICANT: Spriggs, Melanie
APPLICANT: Fanslow, William
TITLE OF INVENTION: No. 6096305e1 Receptor That Binds IL-17
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunex Corporation
STREET: 51 University Street
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Power Macintosh
OPERATING SYSTEM: Apple Operating System 7.5.5
SOFTWARE: Microsoft Word for Apple, Version 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/022,253
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/620,694
FILING DATE: 21-MARCH-1996
APPLICATION NUMBER: USSN 08/538,765
FILING DATE: 7 AUGUST 1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/410,535
FILING DATE: 23 MARCH 1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Perkins, Patricia Anne
REGISTRATION NUMBER: 34,695
REFERENCE/DOCKET NUMBER: 2617-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206)587-0430
```

TELEFAX: (206)  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 158 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-022-253-7

Query Match 13.0%; Score 112; DB 3; Length 158;  
Best Local Similarity 33.3%; Pred. No. 1.7e-05;  
Matches 33; Conservative 14; Mismatches 44; Indels 8; Gaps 5;

QY 32 GGRPADRRRPPPTNL-RSVSPMAYRISYDPARYPRYLPEAYCICRCGCTGLFGEEDVRR 90  
DB 58 GAKVSSR--RPSDYLNRSSTPWLHNRNEDDRPVSIVWEAQCHRCVNA-EKGLDHNM 114

QY 91 SAVVMPYTVLRRTP-ACAGGRSVTEAYVTIPVCTCV 128  
DB 115 SVLIQOEILVLRKRESCP---FTFVEKMLVGVCCTCV 150

## RESULT 9

US-09-022-260-7  
Sequence 7, Application US/09022260  
Patent No. 6100235

## GENERAL INFORMATION:

APPLICANT: Yao, Zhengbin  
APPLICANT: Spriggs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 6100235e1 Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESS: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022.260

## FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/620.694

## FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/410.535  
FILING DATE: 23 MARCH 1995

## CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430

## TELEFAX: (206)

INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 158 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: protein  
US-09-022-260-7

Query Match 13.0%; Score 112; DB 3; Length 158;  
Best Local Similarity 33.3%; Pred. No. 1.7e-05;  
Matches 33; Conservative 14; Mismatches 44; Indels 8; Gaps 5;

QY 32 GGRPADRRRPPPTNL-RSVSPMAYRISYDPARYPRYLPEAYCICRCGCTGLFGEEDVRR 90  
DB 58 GAKVSSR--RPSDYLNRSSTPWLHNRNEDDRPVSIVWEAQCHRCVNA-EKGLDHNM 114

QY 91 SAVVMPYTVLRRTP-ACAGGRSVTEAYVTIPVCTCV 128  
DB 115 SVLIQOEILVLRKRESCP---FTFVEKMLVGVCCTCV 150

## RESULT 10

US-08-620-694A-8  
Sequence 8, Application US/08620694A  
Patent No. 5869286

## GENERAL INFORMATION:

APPLICANT: Yao, Zhengbin  
APPLICANT: Spriggs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 5869286e1 Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESS: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/620.694A

## FILING DATE: 21 MARCH 1996

## CLASSIFICATION: 435

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 08/538,765  
FILING DATE: 7 AUGUST 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/410,535  
FILING DATE: 23 MARCH 1995

## CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430

## TELEFAX: (206)

## INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:  
LENGTH: 151 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO

## ANTI-SENSE: NO

## ORIGINAL SOURCE:

ORGANISM: Herpesvirus Saimiri  
STRAIN: ORF13

US-08-620-694A-8

Query Match 12.6%; Score 108.5; DB 2; Length 151;  
Best Local Similarity 31.4%; Pred. No. 4.1e-05;  
Matches 27; Conservative 11; Mismatches 37; Indels 11; Gaps 3;





GENERAL INFORMATION:  
APPLICANT: Yao, Zhengbin  
APPLICANT: Springs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 6072037e1 Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple Operating System 7.5.5  
SOFTWARE: Microsoft Word for Apple, Version 6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/022,696  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/620,694  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/620,694  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: USSN 08/410,535  
FILING DATE: 23 MARCH 1995  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Perkins, Patricia Anne  
REGISTRATION NUMBER: 34,695  
REFERENCE/DOCKET NUMBER: 2617-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206)587-0430  
TELEFAX: (206)587-0430  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 151 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Herpesvirus Saimiri  
STRAIN: ORF13  
US-09-022-696-8

Query Match 12.6%; Score 108.5; DB 3; Length 151;  
Best Local Similarity 31.4%; Pred. No. 4.1e-05;  
Matches 27; Conservative 11; Mismatches 37; Indels 11; Gaps 3;

QY 47 RSVSPWARYISIDPARYPRLPAYLCRGCLTGLFGEDVRRSAPVYMPVYLRRTPA 106  
DB 65 RSTSPWTLHRNEDQDRYPVIMEAKRCRYLGCYNA-DGNVDYHMSVPIQOELLVVRK--- 120  
QY 107 CAGGRSVYTEAY---VTIPVGCCTV 128  
DB 121 ---GHQPCPNFRLEKMLVTVGCCTV 143

RESULT 14  
US-08-685-239-6  
Sequence 6, Application US/08685239  
Patent No. 6074849  
GENERAL INFORMATION:  
APPLICANT: Jacobs, Kenneth  
APPLICANT: Kelleher, Kerry  
APPLICANT: Carlin, McKeough

APPLICANT: Goldman, Samuel  
APPLICANT: Pittman, Debra  
APPLICANT: M. Sha  
APPLICANT: Neden, Steven  
APPLICANT: Giannotti, JoAnn  
APPLICANT: Golden/Fleet, Margaret  
TITLE OF INVENTION: Human CTLA-8 and Uses of CTLA-8-Related Proteins  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genetics Institute, Inc.  
STREET: 87 CambridgePark Drive  
CITY: Cambridge  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02140  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/685,239  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Brown, Scott A.  
REGISTRATION NUMBER: 32,724  
REFERENCE/DOCKET NUMBER: G15262  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 498-8224  
TELEFAX: (617) 876-5851  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 151 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-685-239-6

Query Match 12.6%; Score 108.5; DB 3; Length 151;  
Best Local Similarity 31.4%; Pred. No. 4.1e-05;  
Matches 27; Conservative 11; Mismatches 37; Indels 11; Gaps 3;

QY 47 RSVSPWARYISIDPARYPRLPAYLCRGCLTGLFGEDVRRSAPVYMPVYLRRTPA 106  
DB 65 RSTSPWTLHRNEDQDRYPVIMEAKRCRYLGCYNA-DGNVDYHMSVPIQOELLVVRK--- 120  
QY 107 CAGGRSVYTEAY---VTIPVGCCTV 128  
DB 121 ---GHQPCPNFRLEKMLVTVGCCTV 143

RESULT 15  
US-09-022-253-8  
Sequence 8, Application US/09022253  
Patent No. 6096305  
GENERAL INFORMATION:  
APPLICANT: Yao, Zhengbin  
APPLICANT: Springs, Melanie  
APPLICANT: Fanslow, William  
TITLE OF INVENTION: No. 6096305e1 Receptor That Binds IL-17  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh

Thu Mar 1 14:24:28 2001

us-09-320-713-4.rai

Page 8

OPERATING SYSTEM: Apple Operating S  
 SOFTWARE: Microsoft Word for Apple,  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/022,253  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/620,694  
 FILING DATE: 21-MARCH-1996  
 APPLICATION NUMBER: USSN 08/358,765  
 FILING DATE: 7 AUGUST 1995  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: USSN 08/410,535  
 FILING DATE: 23 MARCH 1995  
 CLASSIFICATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Perkins, Patricia Anne  
 REGISTRATION NUMBER: 34,695  
 REFERENCE/DOCKET NUMBER: 2617-B  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206)587-0430  
 TELEFAX: (206)  
 INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 151 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: No. 6096305 Relevant  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 ORIGINAL SOURCE:  
 ORGANISM: Herpesvirus Saimiri  
 STRAIN: ORF13  
 US-09-022-253-8

Query Match	12.6%	Score 108.5	DB 3	Length 151
Best Local Similarity	31.4%	Pred. No. 4.1e-05		
Matches 27	Conservative 11	Mismatches 37	Indels 11	Gaps 3
Oy	47	RSVSPMAVAVISDPARPYRPLEPAVCLRGCLTGLFGEDVAFRSAPYVMPVVLRRPA	106	
		:       :         :       :		
Db	65	RSTSPWTLHRNDODDRPSVINEACRIGCVNA-DGNVDYHMNSVPIQOETLIVRK---	120	
Oy	107	CAGGRSVYTYE----VTPVGCSTCV	128	
		: : :		
Db	121	---GHQPCDPSFRLEKMLVTGCTCV	143	

Search completed: February 16, 2001, 12:50:11  
Job time: 131 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2001, 12:50:39 ; Search time 23.57 Seconds  
(without alignments)  
460.929 Million cell updates/sec

Title: US-09-320-713-4

Perfect score: 864  
Sequence: 1 NSARARAVASAFHHTLQLCP.....SIDKQAKLLGPNAPACR 160

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 67900655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 66:\*

1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	113.5	13.1	147	2	JC4628 cytotoxic T-lympho
2	108.5	12.6	151	1	B45351 immediate-early pr
3	108	12.5	150	2	I49623 cytotoxic T-lympho
4	99	11.5	148	2	T21334 hypothetical prote
5	83.5	9.7	226	2	T27843 hypothetical prote
6	77	8.9	354	2	T27099 hypothetical prote
7	77	8.9	603	2	S28941 coagulation factor
8	76	8.8	242	2	B81719 conserved hypotet
9	75.5	8.7	610	2	T44254 thiamin biosynthes
10	75	8.7	178	2	D72281 hypothetical prote
11	75	8.7	242	2	A71568 hypothetical prote
12	75	8.7	362	2	S23395 cyclin-dependent k
13	74.5	8.6	181	2	I68674 cyclin-dependent k
14	74.5	8.6	1273	2	T38292 hypothetical RNA p
15	74.5	8.6	1896	1	RNF21L diaminopimelate de
16	74	8.6	495	2	T09993 versican precursor
17	73	8.4	2397	1	A55535 cyclin-dependent k
18	71.5	8.3	789	2	A39564 transcription repr
19	71.5	8.3	464	2	D72653 hypothetical prote
20	71	8.2	998	2	G83022 probable two-compo
21	70.5	8.2	181	2	I54380 cyclin-dependent k
22	70.5	8.2	256	2	T47860 transcription fact
23	70.5	8.2	265	2	D70778 hypothetical prote
24	70.5	8.2	479	1	A42241 glycine hydroxymet
25	70.5	8.2	872	2	H75564 probable ATP-depen
26	70.5	8.2	2115	2	S38480 nonstructural prot
27	70.5	8.1	206	2	T36643 probable integral
28	70	8.1	206	2	T36643
29	70	8.1	816	2	C69493 hypothetical prote

30	69	8.0	162	2	T32515 hypothetical prote
31	69	8.0	292	2	T03122 hypothetical prote
32	69	8.0	802	2	T24293 hypothetical prote
33	69	8.0	949	2	T24294 hypothetical prote
34	68.5	7.9	262	2	F75093 hypothetical prote
35	68.5	7.9	342	2	A56552 homeotic protein H
36	68	7.9	173	2	E71017 probable NADH-ubiq
37	68	7.9	298	2	A41230 DNA-3-methyladenin-
38	68	7.9	528	2	T00951 probable 3-oxoacyl
39	68	7.9	585	2	A83020 probable cardamom
40	68	7.9	990	2	T14756 hypothetical prote
41	68	7.9	2205	1	MMWVRN nonstructural poly
42	67.5	7.8	298	2	A32872 myogenic factor CM
43	67.5	7.8	298	2	A53783 transcription regu
44	67.5	7.8	426	2	A42360 cellulase (EC 3.2.
45	67.5	7.8	453	2	T01114 hypothetical prote

## ALIGNMENTS

```
RESULT 1
JC4628
cytotoxic T-lymphocyte-associated antigen 8 precursor - mouse
N:Alternate names: CTLA8 protein
C:Species: Mus musculus (house mouse)
C:Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 26-Aug-1999
C:Accession: JC4628
R:Yao, Z.; Timour, M.; Painter, S.; Fanslow, W.; Spriggs, M.
Gene 168, 223-225, 1996
A:Title: Complete nucleotide sequence of the mouse CTLA8 gene.
A:Reference number: JC4628; MUID:96194901
A:Accession: JC4628
A:Molecule type: DNA
A:Residues: 1-147 <YAO>
A:Cross-references: GB:U05108; NID:g1244499; PIDN:AAA93253.1; PID:g1244500
C:Genetics:
A:Gene: ctla8
A:Introns: 69/2
C:Superfamily: salm1rl herpesvirus immediate-early protein 2
C:Keywords: cytokine; glycoprotein; lymphocyte
F:1-14/Domain: signal sequence #status predicted <SIG>
F:15-147/Product: cytotoxic T-lymphocyte-associated antigen 8 #status predicted <MAT>
F:60/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.1%; Score 113.5; DB 2; Length 147;
Best Local Similarity 31.4%; Pred. No. 0.0003;
Matches 37; Conservative 16; Mismatches 46; Indels 19; Gaps 6;

QY 32 GGRPADRRFRPTNL-RSVSPWATRYSTDPARTPLPACVLCRCGLTGLGGEDEVRR 90
      | : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 47 GAKVSR--RPSDYLNRKSTSPWTLHRNEDPDRPSVYWEACRHCRCVNA-BGKLDHNN 103
      | : | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 91 SAPVWPTVLRTRP-ACAGRSVYEAVTIPVCTGCPPEEKDADSINSSIDKGA 147
      | : | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 104 SYLIDQELIVLRPEPSCP---FTFRVEKMLVGVGCTCV-----ASTVRQA 147
      | : | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 2
B45351
immediate-early protein 2 - salm1rl herpesvirus 1 (strain 11)
N:Alternate names: hypothetical protein ORF13
C:Species: salm1rl herpesvirus 1
A:Note: host Salm1rl sclureus (common squirrel monkey)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999
C:Accession: B45351; D36807
R:Nicholas, J.; Smith, E.P.; Coles, L.; Honess, R.
Virology 179, 189-200, 1990
A:Title: Gene expression in cells infected with gammaherpesvirus salm1rl: properties
A:Reference number: A45351; MUID:91021021
A:Accession: B45351
A:Molecule type: mRNA
```



C:Accession: T27099  
 R:Smyle, R.  
 submitted to the EMBL Data Library, October 1998  
 A:Reference number: 220309  
 A:Accession: T27099  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-354 <MIL>  
 A:Cross-references: EMBL:AL032644; PIDs:CAA21670.1; GSPDB:GN00020; CESP:Y51H1A.7  
 A:Experimental source: clone Y51H1A  
 C:Genetics:  
 A:Gene: CESP:Y51H1A.7  
 A:Map position: 2  
 A:Introns: 310/3

Query Match 8.9%; Score 77; DB 2; Length 354;  
 Best Local Similarity 25.0%; Pred. No. 3.5;  
 Matches 36; Conservative 20; Mismatches 56; Indels 32; Gaps 8;

QY 20 PRQANASCPGGRADRRFRPTMLRSVPAYRISIDPARV-PRYLPEATCLRGCL 78  
 DB 141 PERRPRRSAPAS--PAKQYRPQPPARSHEETORIMNKVKEMPRRRRDADYDQ--- 195  
 QY 79 TGLFGEDVFRFSAPYMPVLRRTPCAGRSYTYEAVTIPVGCCTCVPPEKADSI 138  
 DB 196 ----GGQOSRLQ-----MLEQTLRLAP-----LEEEK-----PPVATMAPVDE----- 231  
 QY 139 NSSI--DKQAKLLIGPNDAPAG 160  
 DB 232 HKAIPTWRPGRKRRMGPEMATAP 255

RESULT 7  
 S28941  
 coagulation factor XIIa (BC 3.4.21.38) - guinea pig (fragment)  
 N:Alternate names: Hageman factor  
 C:Species: Cavia porcellus (guinea pig)  
 C:Date: 25-Feb-1994 #sequence\_revision 03-Aug-1995 #text\_change 21-Jan-2000  
 C:Accession: S28941  
 R:Semba, U.; Yamamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kanbara, T.; Okabe, H.  
 Biochim. Biophys. Acta 1159, 113-121, 1992  
 A:Title: Primary structure of guinea-pig Hageman factor: sequence around the cleavage site  
 A:Reference number: S28941; MUID:93003367  
 A:Accession: S28941  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-603 <SEM>  
 A:Cross-references: EMBL:X68615; NID:949578; PIDs:CAA48600.1; PID:949579  
 C:Superfamily: coagulation factor XII; BGF homology; fibrinectin type I repeat homology;  
 C:Keywords: hydrolase; serine proteinase  
 F:46-87/Domain: fibrinectin type II repeat homology <1F>  
 F:134-169/Domain: fibrinectin type I repeat homology <EGF>  
 F:177-208/Domain: EGF homology  
 F:216-294/Domain: kringe homology <KRG>  
 F:359-597/Domain: trypsin homology <TRY>

Query Match 8.9%; Score 77; DB 2; Length 603;  
 Best Local Similarity 21.2%; Pred. No. 6.1;  
 Matches 39; Conservative 18; Mismatches 59; Indels 68; Gaps 8;

QY 12 FHHITLQGFREARNASCPAGRRADRRRPPV-NLRSSVPAYRISIDPARVPRY---- 66  
 DB 57 YHH-----CINHGKRPGRPCATTPNFDODQAVYCL--EPKKYKDCSKH 100  
 QY 67 -----LPEAYCCRCGLTG-----LFGEDVFRFSAPV----- 94  
 DB 101 NPGQRGICVNTLSSPHCLCPDLHLTKKQOREKCFERQLHRRFHEHIEIFRTGPAVYAKC 160  
 QY 95 -----YPTVVLARTTPACAGRSYTYEAV--VTIPVGCCTCVPPEKADSI 142  
 DB 161 HCKGPDANCKOMHSRCQTNPCINGRCLEVEGHNLCDCRPMGYT--GPFCDLDTTASCY 217

QY 143 DKOG 146  
 DB 218 EGRG 221

RESULT 8  
 B81719  
 conserved hypothetical protein TC0284 [imported] - Chlamydia muridarum (strain Nigg)  
 C:Species: Chlamydia muridarum, Chlamydia trachomatis Mopn  
 C:Date: 31-Mar-2000 #sequence\_revision 31-Mar-2000 #text\_change 11-May-2000  
 C:Accession: B81719  
 R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke,  
 C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,  
 Nucleic Acids Res. 28, 1397-1406, 2000  
 A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39  
 A:Reference number: A81500; MUID:20150255  
 A:Accession: B81719  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-242 <TEF>  
 A:Cross-references: GB:AE002296; GB:AE002160; NID:g7190325; PIDs:AAF39152.1; PID:g719  
 A:Experimental source: strain Nigg (Mopn)  
 C:Genetics:  
 A:Gene: TC0284

Query Match 8.8%; Score 76; DB 2; Length 242;  
 Best Local Similarity 26.2%; Pred. No. 2.9;  
 Matches 21; Conservative 6; Mismatches 21; Indels 32; Gaps 3;

QY 68 PEAYCICRCGLTGLF-----GEQVFRFSAPYMPVLRRTPCAGRSYTY---- 115  
 DB 161 PEPHCNCLHCQIGRIAVEEDIEVSEDLTFRSMWI-----SQYGERMYTVTD 208  
 QY 116 -----EAYVTRPYGCTC 127  
 DB 209 PLNPEQGFNVYLGTPIGCTC 228

RESULT 9  
 T44254  
 thiamin biosynthesis protein thic [imported] - Rhizobium etli plasmid b  
 C:Species: Rhizobium etli  
 C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jul-2000  
 C:Accession: T44254  
 R:Miranda-Rios, J.; Moreira, C.; Taboada, H.; Davalos, A.; Encarnacion, S.; Mora, J.;  
 J. Bacteriol. 179, 6887-6893, 1997  
 A:Title: Expression of thiamin biosynthetic genes (thiCOGE) and production of symbiot  
 A:Reference number: 222737; MUID:98037482  
 A:Accession: T44254  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-610 <MIR>  
 A:Cross-references: EMBL:AF004408; NID:g2627325; PIDs:AMC45972.1; PID:g2627326  
 A:Experimental source: strain CE3  
 C:Genetics:  
 A:Gene: thic  
 A:Genome: plasmid b  
 C:Function:  
 A:Description: involved in the biosynthesis of the pyrimidine moiety of thiamin  
 C:Superfamily: thiamin biosynthesis protein thic

Query Match 8.7%; Score 75.5; DB 2; Length 610;  
 Best Local Similarity 29.9%; Pred. No. 8.8;  
 Matches 38; Conservative 14; Mismatches 54; Indels 21; Gaps 7;

QY 30 PACGRPADRRFRPTMLRSVPW---AYRISIDPARVPRYLPEAYCICRCGLTGLFGE-- 84  
 DB 44 PTSGEP-----PVTVYDSGSPYTDPAHVISID-AGLPR-LRESWIKARGDVESYDGRIV 95  
 QY 85 --EDVAFRSAPYMPVLRRTPCA-GGRSYTYEAVTIPVGCCTCVPPEKADSI 141



Db 49 -----RDCDALMACGICQIQAERHNFVETPLEGDFAMERVGLGLPKLYLPTGPRGR 103  
 QY 105 PACAGRSYVT-----EAYTIPGCTCVPPEPENDAD-SINSSIDKQKAK 148  
 Db 104 DELGGRRGRTSPALLQGTAEEDHVDLSICTLVPRSGQAGSPGGPSGGRK 158

# RESULT 14

738292

hypothetical protein SPAC23E2.02 - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T38292

R:Skellon, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Walsby, S.V.

Submitted to the EMBL Data Library, January 1996

A:Reference number: 221764

A:Accession: T38292

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Cross-references: EMBL:Z68887; PIDN:CAA93114.1; GSPDB:GN00066; SPDB:SPAC23E2.02

C:Genetics:

A:Gene: SPDB:SPAC23E2.02

A:Map position: 1

A:Introns: 8/2; 862/3

Query Match 8.6%; Score 74.5; DB 2; Length 1273;

Best Local Similarity 25.6%; Pred. No. 24;

Matches 22; Conservative 15; Mismatches 48; Indels 1; Gaps 1;

QY 73 ICRGLTGLFGEEDVFRSAPVYMPVLRTPACAGSVTEAVTIPVCTCVPPE 132  
 Db 520 ICAQLTGLFGEEDVFRSAPVYMPVLRTPACAGSVTEAVTIPVCTCVPPE 132  
 QY 133 KDADSISSIDKQAKLLGPNDA 158  
 Db 579 SNSNSISNSTSLNPKRDVDPSPHS 604

# RESULT 15

RNF2L

DNA-directed RNA polymerase (EC 2.7.7.6) II 215K chain [validated] - fruit fly (Drosophila)

C:Species: Drosophila melanogaster

C:Date: 04-Dec-1986 #sequence\_revision 31-Mar-1993 #text\_change 28-Jul-2000

C:Accession: S04457; A00693; B27677; S60151

R:Jokierst, R.S.; Weeks, J.R.; Zehring, W.A.; Greenleaf, A.L.

Mol. Gen. Genet. 215, 266-275, 1989

A:Title: Analysis of the gene encoding the largest subunit of RNA polymerase II in Drosophila

A:Reference number: S04457; MUID:89218930

A:Accession: S04457

A:Molecule type: DNA

A:Cross-references: EMBL:M27431; NID:q158331; PIDN:AAA28868.1; PID:q158332

R:Biggs, J.; Seales, L.L.; Greenleaf, A.L.

Cell 42, 611-621, 1985

A:Title: Structure of the eukaryotic transcription apparatus: features of the gene for the largest subunit of RNA polymerase II

A:Reference number: A00693; MUID:85282618

A:Accession: A00693

A:Molecule type: DNA

A:Residues: 1-318, 'GVAKV', 325-449, 'G', 451-454, 'RCTT', 459-462, 'VTGESVASS', <BIG>

A:Cross-references: EMBL:M11798

A>Note: This sequence has been revised in reference S04457

R:Allison, L.A.; Wong, J.K.C.; Fitzpatrick, V.D.; Moyle, M.; Ingles, C.J.

Mol. Cell. Biol. 8, 321-329, 1988

A:Title: The C-terminal domain of the largest subunit of RNA polymerase II of Saccharomyces

A:Reference number: A93104; MUID:88094402

A:Accession: B27677

A:Molecule type: DNA

A:Residues: 1441-1484, 'I', 1527-1889 <ALL>

A:Cross-references: EMBL:M19537; NID:q158147; PIDN:AAA28827.1; PID:q158148

R:Peterson, G.; Song, D.; Huegle-Doerr, B.; Oldenburg, I.; Bantz, E.K.F.

Mol. Gen. Genet. 249, 425-431, 1995

A:Title: Mapping of linear epitopes recognized by monoclonal antibodies with gene-fra

A:Reference number: S60151; MUID:96133682

A:Accession: S60151

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 778-827 <PET>

C:Genetics:

A:Gene: RPO21; Rpi1215

A:Cross-references: FlyBase:FBgn0003277

A:Map position: X 10C, X 35.7

A:Introns: 27/3; 775/3; 1526/1

C:Function:

A:Description: EC 2.7.7.6 [validated; MUID:88094402]; essential for proper initiation

C:Superfamily: human DNA-directed RNA polymerase II largest chain

C:Keywords: DNA binding; nucleotidyltransferase; tandem repeat; transcription; zinc f

F:67-83/Region: zinc finger CCH motif

F:1581-1883/Region: 7-residue repeats

F:349/Binding site: ATP/GTP (Lys) #status predicted

Query Match 8.6%; Score 74.5; DB 1; Length 1896;

Best Local Similarity 24.8%; Pred. No. 37;

Matches 37; Conservative 15; Mismatches 50; Indels 47; Gaps 9;

QY 20 PREQARNASCPRAGRPDRFRPPT-NLRVSPMAVRIYDARY-----PRIPEAYCIC 74  
 Db 1583 PTPSPNYTASSPGGASP-----NYSFSSPNYSPTSP-----LYASPRVASTPNNPOS----- 1630  
 QY 75 RGLTGLFGEEDVFRSAPVYMPVLRTPACAG-GRSVYT--EAYVTIPGCTCVPPE 132  
 Db 1631 ----TGYSSSSGYSPTSPVYPTVOFQSSPSRAGSGSNITSPGNAY-----SP 1675  
 QY 132 EKDADSISSIDKQAKLLGPNDA 160  
 Db 1676 SSSNYSNPS-----PSYSPSP 1692

Search completed: February 16, 2001, 12:50:42

Job time: 148 sec





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 16, 2001, 12:52:57 ; Search time 14.73 Seconds  
(without alignments)  
350.784 Million cell updates/sec

Title: US-09-320-713-4  
Perfect score: 864  
Sequence: 1 NSARARAVLSAFHHTQLGP.....SIDKQAKLLGPNDAPAGP 160

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 88757 seqs, 32294092 residues  
Total number of hits satisfying chosen parameters: 88757

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_39:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	135.5	15.7	155	1 IL17_HUMAN	Q16552 homo sapien
2	113.5	13.1	158	1 IL17_MOUSE	Q62386 mus musculu
3	108.5	12.6	151	1 VG13_HSVSA	P24916 herpesvirus
4	108	12.5	150	1 IL17_RAT	Q61453 rattus norv
5	78.5	9.1	1075	1 NRC3_MOUSE	P97305 mus musculu
6	77	8.9	603	1 FA12_CAVPO	Q04962 cavia porce
7	75.5	8.7	610	1 TH1C_RH1ET	Q34291 rhizobium e
8	75	8.6	362	1 A2HS_PIG	P29700 sus scrofa
9	74.5	8.6	1273	1 VAR2_SCHPO	Q10135 schizosach
10	74.5	8.6	1896	1 RBP1_DROME	P04052 drosophila
11	73.5	8.5	1033	1 CR2_HUMAN	P20023 homo sapien
12	73	8.4	3358	1 PCGV_MOUSE	Q62059 mus musculu
13	71.5	8.3	164	1 CDN1_HUMAN	P38936 homo sapien
14	70.5	8.2	265	1 YM40_MYCTU	Q10522 mycobacteri
15	70.5	8.2	479	1 GLYC_NEUCR	P34898 neurospora
16	69.5	8.0	324	1 TTP_BOVIN	P53781 bos taurus
17	68.5	7.9	342	1 HXCA_MOUSE	P31257 mus musculu
18	68	7.9	2205	1 POLN_RUBVT	P13889 rubella vir
19	67.5	7.8	298	1 MYOD_CHICK	P16075 gallus gall
20	67.5	7.8	384	1 DUS9_HUMAN	Q09956 homo sapien
21	67.5	7.8	591	1 UR49_HBV	P14347 Epstein-Bar
22	67	7.8	277	1 HXDB_MOTVI	P31263 neoptulham
23	67	7.8	2273	1 HPA1_YEAST	P32874 saccharomyc
24	67	7.8	4393	1 PGBM_YEAST	P18160 homo sapien
25	66.5	7.7	2339	1 CCAB_RABIT	Q00152 oryctolagus
26	66	7.6	489	1 VHS_HSV1	P10225 herpes simp
27	66	7.6	676	1 DNLJ_THERH	P26996 thermus aqu
28	65.5	7.6	170	1 PLGF_HUMAN	P48763 homo sapien
29	65.5	7.6	369	1 CHRB_RHOSH	Q33558 rhododactery
30	65.5	7.6	404	1 VE2_HPV60	Q08094 human papil
31	65.5	7.6	500	1 STCL_EMBNI	Q00707 emericella
32	65	7.5	345	1 YJUN_ECOLI	P39400 escherichia
33	65	7.5	376	1 METK_AQUAE	O67222 aquiflex aeo

34	65	7.5	802	1	ACSB_ACEXY	P37716 acetobacter
35	65	7.5	1106	1	GLI1_HUMAN	P08151 homo sapien
36	64.5	7.5	285	1	HXA4_MOUSE	P06798 mus musculu
37	64.5	7.5	374	1	RLIG_BPT4	P00971 bacterioph
38	64	7.4	286	1	PYR1_SYNEL	P50034 synecchococ
39	64	7.4	453	1	CARA_NEUCR	P22572 neurospora
40	64	7.4	497	1	CPDH_MACFA	Q29488 macaca fasc
41	64	7.4	3396	1	PCGV_HUMAN	P13611 homo sapien
42	63.5	7.3	258	1	U149_HSVBC	P30022 bovine hept
43	63.5	7.3	419	1	IRF3_MOUSE	P70671 mus musculu
44	63.5	7.3	429	1	ELK1_MOUSE	P41969 mus musculu
45	63.5	7.3	476	1	CB55_PAPGL	Q95036 papilio gla

## ALIGNMENTS

RESULT	1	IL17_HUMAN	STANDARD:	PRT:	155 AA.
AC	Q16552:	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)				
DT	15-DEC-1998 (Rel. 37, Last annotation update)				
DE	INTERLEUKIN-17 PRECURSOR (IL-17) (CYTOTOXIC T LYMPHOCYTE-ASSOCIATED ANTIGEN 8) (CTLA-8).				
DE	IL17 OR CTLA8.				
GN	Homo sapiens (Human).				
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OC	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=96281911; PubMed=8676080;				
RA	Fossiez F., Djossou O., Chomarat P., Flores-Romo L., Alt-Yahia S., Maat C., Pin J.-D., Garrone P., Garcia E., Saezland S., Blanchard D., Gaillard C., Das Mahapatra B., Rouvier E., Golstein P., Banchereau J., Lebecque S.;				
RA	"T cell interleukin-17 induces stromal cells to produce proinflammatory and hematopoietic cytokines.";				
RT	J. Exp. Med. 183:2593-2603(1996).				
RL	[2]				
RN	SEQUENCE FROM N.A.				
RP	TISSUE=T-CELL;				
RC	MEDLINE=96094436; PubMed=7499828;				
RX	Yao Z., Palanter S.L., Fanslow W.C., Ulrich D., Macduff B.M., Spriggs M.K., Arnltage R.J.;				
RA	"Human IL-17: a novel cytokine derived from T cells.";				
RT	J. Immunol. 155:5483-5486(1995).				
RL	J. Immunol. 155:5483-5486(1995).				
CC	- FUNCTION: INDUCES STROMAL CELLS TO PRODUCE PROINFLAMMATORY AND HEMATOPOIETIC CYTOKINES. ENHANCES THE SURFACE EXPRESSION OF THE INTRACELLULAR ADHESION MOLECULE-1 (ICAM-1) IN FIBROBLASTS.				
CC	- SUBUNIT: HOMODIMER.				
CC	- SUBCELLULAR LOCATION: SECRETED.				
CC	- TISSUE SPECIFICITY: RESTRICTED TO ACTIVATED MEMORY T-CELLS.				
CC	- PFM: FOUND BOTH IN GLYCOSYLATED AND NONGLYCOSYLATED FORMS.				
CC	- SIMILARITY: STRONG, TO OTHER MAMMALIAN INTERLEUKIN-17 AND TO HERPESVIRUS SAIMIRI IMMEDIATE EARLY GENE 13 PROTEIN.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).				
CC	-----				
DR	EMBL: Z58820; CA91233.1; -				
DR	EMBL: U32659; AAC50341.1; -				
DR	MDM: 603149; -				
KW	Cytokine; Glycoprotein; T-cell; Antigen; Signal.				
FT	SIGNAL	1	23	POTENTIAL.	
FT	CHAIN	24	155	INTERLEUKIN-17.	
FT	CARBOHD	68	68	N-LINKED (GLCNAC. . . ) (POTENTIAL).	





CC -1- DOMAIN: REL SIMILARITY DOMAIN (RSD) ALLOWS DNA-BINDING AND  
 CC COOPERATIVE INTERACTIONS WITH API FACTORS (BY SIMILARITY).  
 CC -1- PTM: PHOSPHORYLATED BY NEATC-KINASE; DEPHOSPHORYLATED BY  
 CC CALCIUMKININ (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D85612; BAA12833.1; -  
 DR EMBL; U28807; AAA93249.1; -  
 DR MGD; MG1.103296; NFATC3.  
 DR INTERPRO; IPR000451; -  
 DR PROSITE; PS01204; REL\_1; FALSE-NEG.  
 DR PROSITE; PS01254; REL\_2; 1.  
 KM Transcription regulation; Activator; Nuclear protein; DNA-binding;  
 KM Alternative splicing; Repeat; Phosphorylation.  
 FT DOMAIN 24 29  
 FT 109 114  
 FT 207 308  
 FT 207 223  
 FT 236 252  
 FT 292 308  
 FT 273 275  
 FT 444 451  
 FT 686 688  
 FT 1031 1040  
 FT 468 497  
 FT 1035 1075  
 FT VARSPPLIC 1035 1075  
 FT VARSPPLIC 1035 1075  
 FT CONFLICT 12 34  
 FT 57 61  
 FT 68 83  
 FT 82 83  
 FT 90 93  
 FT 113 113  
 FT 121 124  
 FT 132 132  
 FT 140 140  
 FT 641 641  
 FT 646 646  
 FT 707 734  
 FT CONFLICT 746 746  
 FT 1075 1075  
 FT 1075 1075  
 SO SEQUENCE 1075 AA; 115450 MW; 4ED38C9AAGF452BB CRC64;  
 Query Match 9.18; Score 78.5; DB 1; Length 1075;  
 Best Local Similarity 22.08; Pred. No. 2.8;  
 Matches 37; Conservative 22; Mismatches 76; Indels 33; Gaps 5;  
 1 NSARARAVLSAFHHTLQGPREGQARNASCPA-----GGRPADRRRRP 42  
 684 NGKRRKSSQSRFTYTPVLMKQEDREDTLPSPVSLPVPBSAQORPSSETHPHDRAMSA 743  
 43 PTNLR-SVSPNAVYRISYDPAARYPRYLPAAYCLRGCLGLGEEDVRRSAPVYMPVYL 101  
 744 PGLGLQVOP-----AYISMAVASHLPOLQCRDEGA-----GKEQHATSSVMHOPFOVT 793  
 102 RRTPACAGRSVTEAYVTIPVGTCTVPEPRKADSISSIDKOGAKL 149  
 794 PTPISGSSYSQISQTSMT-----NGPTCLPVPNVASSQSEFVPLEQDAAL 837

FA12\_CAVPO  
 ID FA12\_CAVPO STANDARD; PRT; 603 AA.  
 AC 004962;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE COAGULATION FACTOR XII PRECURSOR (EC 3.4.21.38) (HAGEMAN FACTOR)  
 DE (HAF) (FRAGMENT).  
 GN F12.  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-37, 318-332 AND 359-373.  
 RC TISSUE=LIVER.  
 RX MEDLINE=93003367; Pubmed=1390917;  
 RA Samba U., Yamamoto T., Kunisada T., Shibuya Y., Tanase S.,  
 RA Kambara T., Okabe H.;  
 RT "Primary structure of guinea-pig Hageman factor: sequence around the  
 RT cleavage site differs from the human molecule."  
 RL Biochim. Biophys. Acta 1159:113-121(1992).  
 CC -1- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN  
 CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE  
 CC GENERATION OF BRADYKININ AND ANGIOTENSIN.  
 CC -1- CATALYTIC ACTIVITY: CLEAVES SELECTIVELY ARG-I-ILE BONDS AND  
 CC ACTIVATES COAGULATION FACTORS VII AND XI.  
 CC -1- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A  
 CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY  
 CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST  
 CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR  
 CC XIIA ACTIVATES FACTOR XI TO FACTOR XIA.  
 CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE-I DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE-II DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 1 KRINGLE REGION.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X68615; CAA48600.1; -  
 DR HSPB; P00763; IDPO.  
 DR MEROPS; S01.211; -  
 DR INTERPRO; IPR000001; -  
 DR INTERPRO; IPR000083; -  
 DR INTERPRO; IPR000561; -  
 DR INTERPRO; IPR000562; -  
 DR INTERPRO; IPR001254; -  
 DR PFAM; PF00008; EGF\_2.  
 DR PFAM; PF00039; fn1; 1.  
 DR PFAM; PF00040; fn2; 1.  
 DR PFAM; PF00051; kringle; 1.  
 DR PFAM; PF00089; trypsin; 1.  
 DR PROSITE; PS00021; KRINGLE\_1; 1.  
 DR PROSITE; PS00022; EGF\_1; 2.  
 DR PROSITE; PS00023; FIBRONECTIN\_2; 1.  
 DR PROSITE; PS00134; TRYPSIN\_HIS\_1; 1.  
 DR PROSITE; PS00135; TRYPSIN\_SER; 1.  
 DR PROSITE; PS0186; EGF\_2; 1.  
 DR PROSITE; PS01253; FIBRONECTIN\_1; 1.  
 DR PROSITE; PS50070; KRINGLE\_2; 1.  
 DR Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;  
 KM Hydrolyase; Fibrinolysis; EGF-like domain; Repeat; zymogen; Signal.  
 FT SIGNAL 1 18  
 FT CHAIN 19 358  
 FT CHAIN 359 603  
 FT ALPHA-FACTOR XIIA HEAVY CHAIN.  
 FT ALPHA-FACTOR XIIA LIGHT CHAIN.

```

FT DOMAIN 46 87 FIBRONECTIN TYPE-II.
FT DOMAIN 130 130 EGF-LIKE 1.
FT DOMAIN 132 172 FIBRONECTIN TYPE-I.
FT DOMAIN 173 209 EGF-LIKE 2.
FT DOMAIN 216 294 KRINGLE.
FT DOMAIN 312 342 PRO-RICH.
FT DOMAIN 359 603 CATALYTIC.
FT ACT_SITE 398 398 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 447 447 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 551 551 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 551 551 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 97 109 BY SIMILARITY.
FT DISULFID 103 118 BY SIMILARITY.
FT DISULFID 120 129 BY SIMILARITY.
FT DISULFID 134 162 BY SIMILARITY.
FT DISULFID 160 169 BY SIMILARITY.
FT DISULFID 177 188 BY SIMILARITY.
FT DISULFID 182 197 BY SIMILARITY.
FT DISULFID 199 208 BY SIMILARITY.
FT DISULFID 216 294 BY SIMILARITY.
FT DISULFID 237 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 345 472 BY SIMILARITY.
FT DISULFID 383 399 BY SIMILARITY.
FT DISULFID 391 461 BY SIMILARITY.
FT DISULFID 422 425 BY SIMILARITY.
FT DISULFID 488 557 BY SIMILARITY.
FT DISULFID 520 536 BY SIMILARITY.
FT DISULFID 547 578 BY SIMILARITY.
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 603 AA; 66795 MW; 480C6B946FB9ED59 CRC64;

```

```

Query Match 8.9%; Score 77; DB 1; Length 603;
Best Local Similarity 21.2%; Pred. No. 2.1;
Matches 39; Conservative 18; Mismatches 59; Indels 68; Gaps 8;

```

```

QY 12 FHHTDLOGPREQARNASCPAGGRPPRRPPR-NLRSPVARYRISYDAPARYPRY-----66
DB 57 YHH-----CJHKGRRPGRPCWATTPNFDQDMAYCL--EPKKVKYKDCSKH 100
QY 67 -----LPEAYCCRGCLTG-----LFGEDVFRSAPV-----94
DB 101 NPGRGICVNTLSSPHCLCPDHLTGKHCOREKCFEPQLHFRFHEHINFRTPAGVAKC 160
QY 95 -----YMPVTVLRTPACAGGRSVYTEAY--VTIPVGCTCVPPEPKADSIINSI 142
DB 161 HCKGPDHCKOMHSQECQTFNPLCNGRCLEVEGHILCDECMGYT---GPPCDLDTTASCY 217
QY 143 DKOG 146
DB 218 EGRG 221

```

RESULT 7  
THIC\_RHIEF  
ID THIC\_RHIEF STANDARD; PRT; 610 AA.  
AC 034291;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE THIAMINE BIOSYNTHESIS PROTEIN THIC.  
GN THIC.  
OS Rhizobium etl.  
OG Rhizobium etl.  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CE3;  
RX MEDLINE-98037482; PubMed-9371431;  
RA Miranda-Rios J., Morena C., Taboada H., Davalos A., Encarnacion S.,

```

RA Mora J., Soberon M.;
RT "Expression of thiamine biosynthetic genes (thiCGE) and production of
RT symbiotic terminal oxidase cbb3 in Rhizobium etl.";
RL J. Bacteriol. 179:6867-6893(1997).
CC -1- FUNCTION: REQUIRED FOR THE SYNTHESIS OF THE HYDROMETHYLPYRIMIDINE
CC (HMP) MOIETY OF THIAMINE (4-AMINO-2-METHYL-5-
CC HYDROXYMETHYLPYRIMIDINE) (BY SIMILARITY).
CC -1- PATHWAY: THIAMINE BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE THIC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF004408; AAC45972.1; -.
DR INTERPRO: IP002817; -.
DR PRAM: PF01964; Thic; 1.
KW Thiamine biosynthesis; Plasmid.
SQ SEQUENCE 610 AA; 67105 MW; 9CE7F560DB35ACDC CRC64;

```

```

Query Match 8.7%; Score 75.5; DB 1; Length 610;
Best Local Similarity 29.9%; Pred. No. 3;
Matches 38; Conservative 14; Mismatches 54; Indels 21; Gaps 7;

```

```

QY 30 PAGGRADRRFRPPNLRSPV---AYRISYPARYPRYLPRAYCLRCGLTGLGE--284
DB 44 PTSGEP-----PVYYDSSGPTTDAHAHYSID-AGLPR-LRSWIKAKGDVSYGRIV 95
QY 85 --EDVFRSAPVYMPVTVLRTPACA--GGRSVYTEAYVTIPVGCTCVPPEPKADSIINSI 141
DB 96 KPEDNCFATGGERLTPREPVRNPLKAKAGRAYQLAVAR-----AGIVPEMEFIAREK 150
QY 142 IDKOGAK 148
DB 151 LGRQAK 157

```

```

RESULT 8
A2HS_PIG 8
ID A2HS_PIG STANDARD; PRT; 362 AA.
AC P29700;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE ALPHA-2-HS-GLYCOPROTEIN PRECURSOR (FETUIN) (FRAGMENT).
GN AHSg.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE-LIVER;
RX MEDLINE-92209519; PubMed-1372866;
RA Brown W.M., Christie D.L., Saunders N.R., Nawratil P.,
RA Dziegielewska K.D., Mueller-Esterl W.;
RT "The nucleotide and deduced amino acid structures of sheep and pig
RT fetuin. Common structural features of the mammalian fetuin family";
RL Eur. J. Biochem. 203:321-331(1992).
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE FETUIN FAMILY.
CC -1- SIMILARITY: CONTAINS 2 CYSTEINE-LIKE REPEATS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```





FT DISULFID 242 271 BY SIMILARITY.  
 FT DISULFID 276 325 BY SIMILARITY.  
 FT DISULFID 305 342 BY SIMILARITY.  
 FT DISULFID 351 393 BY SIMILARITY.  
 FT DISULFID 379 406 BY SIMILARITY.  
 FT DISULFID 410 453 BY SIMILARITY.  
 FT DISULFID 439 466 BY SIMILARITY.  
 FT DISULFID 471 509 BY SIMILARITY.  
 FT DISULFID 495 522 BY SIMILARITY.  
 FT DISULFID 527 576 BY SIMILARITY.  
 FT DISULFID 556 593 BY SIMILARITY.  
 FT DISULFID 602 644 BY SIMILARITY.  
 FT DISULFID 630 657 BY SIMILARITY.  
 FT DISULFID 662 699 BY SIMILARITY.  
 FT DISULFID 685 714 BY SIMILARITY.  
 FT DISULFID 719 762 BY SIMILARITY.  
 FT DISULFID 748 779 BY SIMILARITY.  
 FT DISULFID 788 830 BY SIMILARITY.  
 FT DISULFID 816 843 BY SIMILARITY.  
 FT DISULFID 851 894 BY SIMILARITY.  
 FT DISULFID 880 907 BY SIMILARITY.  
 FT DISULFID 912 955 BY SIMILARITY.  
 FT DISULFID 941 968 BY SIMILARITY.  
 FT CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 492 492 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 682 682 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 800 800 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 823 823 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 861 861 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 911 911 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 667 667 Q -> G (IN REF. 2).  
 FT CONFLICT 902 902 H -> L (IN REF. 2).  
 FT CONFLICT 906 906  
 SQ SEQUENCE 1033 AA: 112973 MW: 1749DB4A07847ADA CRC64;

Query Match 8.5%: Score 73.5; DB 1; Length 1033;  
 Best Local Similarity 25.0%: Pred. No. 8.8;  
 Matches 32; Conservative 8; Mismatches 41; Indels 47; Gaps 5;  
 QY 42 PPTNLRSVSPWARYLSYDPAARYLPAYCLRCGLTG--LFGEDVFRSAPVYMPV 99  
 DB 413 PPNILNGCKDRHWRPDPGSIKY-----SCNPGYLVGESIQCTSEGVWTFPV 463  
 QY 100 VLKRPAC-AGR-----SVYEAAYTIP-----V 123  
 DB 464 PCKVAACEATGRQLTRPOHQFVVRPDVNSCGEGYKLSGVYQCCGTIPFMEIRLCK 523  
 QY 124 GCTCVPEP 131  
 DB 524 EITCPEPP 531  
 RESULT 12  
 PGCV\_MOUSE STANDARD: PRT: 3358 AA.  
 ID PGCV\_MOUSE  
 AC 062059; 062058;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE VERSION CORE PROTEIN PRECURSOR (LARGE FIBROBLAST PROTEOGLYCAN)  
 DE (CHONDROITIN SULFATE PROTEOGLYCAN CORE PROTEIN 2) (PG-M).  
 GN CSFG2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 RN [1]  
 RP SEQUENCE FROM N.A. (VARIANTS V0; V1 AND V2).  
 RC STRAIN=C57BL/6, AND SWISS WEBSTER; TISSUE=BRAIN;

RX MEDLINE=95122551; PubMed=7822336;  
 RA Ito K., Shinomura T., Zako M., Ujita M., Kimata K.;  
 RT "Multiple forms of mouse PG-M, a large chondroitin sulfate  
 RT proteoglycan generated by alternative splicing.";  
 RT J. Biol. Chem. 270:958-965(1995).  
 RN [2]  
 RP SEQUENCE OF 1-348 AND 3053-3358 FROM N.A. (VARIANT V3).  
 RC STRAIN=C57BL/6;  
 RX MEDLINE=95181355; PubMed=7876137;  
 RA Zako M., Shinomura T., Ujita M., Ito K., Kimata K.;  
 RT "Expression of PG-M(V3), an alternatively spliced form of PG-M  
 RT without a chondroitin sulfate attachment in region in mouse and human  
 RT tissues.";  
 RL J. Biol. Chem. 270:3914-3918(1995).  
 CC -1- FUNCTION: MAY PLAY A ROLE IN INTERCELLULAR SIGNALING AND IN  
 CC CONNECTING CELLS WITH THE EXTRACELLULAR MATRIX. MAY TAKE PART IN  
 CC THE REGULATION OF CELL MOTILITY, GROWTH AND DIFFERENTIATION. BINDS  
 CC HYALURONIC ACID.  
 CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX.  
 CC -1- ALTERNATIVE PRODUCTS: AT LEAST 4 ISOFORMS; V0 (SHOWN HERE), V1,  
 CC V2 AND V3; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: V2 IS FOUND ONLY IN BRAIN.  
 CC -1- DEVELOPMENTAL STAGE: DISAPPEARS AFTER THE CARTILAGE DEVELOPMENT.  
 CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 SUSHI (SCR) REPEAT.  
 CC -1- SIMILARITY: BELONGS TO THE AGGRECAN/VERSICAN PROTEOGLYCAN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D16263; BAA03796.1; -;  
 DR EMBL; D28599; -; NOT ANNOTATED\_CDS.  
 DR EMBL; D32040; BAA06802.1; -;  
 DR HSSP; P00740; ITXA.  
 DR MGD; MGI:102889; CSFG2.  
 DR INTERPRO; IPR000152; -;  
 DR INTERPRO; IPR000436; -;  
 DR INTERPRO; IPR000538; -;  
 DR INTERPRO; IPR000561; -;  
 DR INTERPRO; IPR001304; -;  
 DR INTERPRO; IPR001438; -;  
 DR INTERPRO; IPR001881; -;  
 DR INTERPRO; IPR003006; -;  
 DR PFAM; PF00008; EGF\_2.  
 DR PFAM; PF00193; Xlink; 2.  
 DR PFAM; PF00047; Ig; 1.  
 DR PFAM; PF00059; lectin\_C; 1.  
 DR PFAM; PF00084; sush1; 1.  
 DR PRINTS; PR00010; EGBLOOD.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 1.  
 DR PROSITE; PS00022; EGF\_1; 2.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS01187; EGF\_CA; 1.  
 DR PROSITE; PS01241; LINK; 2.  
 DR PROSITE; PS00615; C-TYPE LECTIN\_1; 1.  
 DR PROSITE; PS00041; C-TYPE LECTIN\_2; 1.  
 KW Glycoprotein; Proteoglycan; lectin; Extracellular matrix; Sush1;  
 KW Signal; Repeat; EGF-like domain; Immunoglobulin domain;  
 KW Hyaluronic acid; Alternative splicing.  
 FT SIGNAL 1 20  
 FT CHAIN 21 3358  
 FT DOMAIN 37 137  
 FT DOMAIN 167 244  
 FT DOMAIN 265 334  
 FT DOMAIN 348 1308  
 GAG-ALPHA (GLUCOSAMINOGLYCAN ATTACHMENT



FT DOMAIN 1309 3052 DOMAIN).  
 FT DOMAIN 3052 3088 EGF-BETA.  
 FT DOMAIN 3090 3126 EGF-LIKE 1.  
 FT DOMAIN 3129 3255 EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).  
 FT REPEAT 3258 3316 C-TYPE LECTIN.  
 FT DISULFID 44 130 SUSHI.  
 FT DISULFID 172 243 BY SIMILARITY.  
 FT DISULFID 196 217 BY SIMILARITY.  
 FT DISULFID 270 333 BY SIMILARITY.  
 FT DISULFID 294 315 BY SIMILARITY.  
 FT DISULFID 3056 3067 BY SIMILARITY.  
 FT DISULFID 3061 3076 BY SIMILARITY.  
 FT DISULFID 3078 3087 BY SIMILARITY.  
 FT DISULFID 3094 3105 BY SIMILARITY.  
 FT DISULFID 3099 3114 BY SIMILARITY.  
 FT DISULFID 3116 3125 BY SIMILARITY.  
 FT DISULFID 3132 3143 BY SIMILARITY.  
 FT DISULFID 3160 3252 BY SIMILARITY.  
 FT DISULFID 3228 3244 BY SIMILARITY.  
 FT DISULFID 3259 3302 BY SIMILARITY.  
 FT DISULFID 3288 3315 BY SIMILARITY.  
 FT CARBOHYD 57 57 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 807 807 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 914 914 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 951 951 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1305 1305 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1372 1372 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1679 1679 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 2054 2054 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 2244 2244 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 2362 2362 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 2627 2627 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 3030 3030 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 3332 3332 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 3342 3342 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 349 1308 MISSING (IN ISOFORM V1).  
 FT VARSPLIC 1309 3052 MISSING (IN ISOFORM V2).  
 FT VARSPLIC 349 3052 MISSING (IN ISOFORM V3).  
 FT CONFLICT 348 348 P -> R (IN REF. 2).  
 SQ SEQUENCE 3358 AA; 366938 MW; 071B80026BC0762D CRC64;

Query Match 8.4%; Score 73; DB 1; Length 3358;  
 Best Local Similarity 32.1%; Pred. No. 38;  
 Matches 26; Conservative 10; Mismatches 33; Indels 12; Gaps 4;

QY 81 LGGEYVRRSARVYMPYVLRTPACAGRSYV-TE-AVYTPVCGCTCP-----EPEK 133  
 DB 3035 LGISESEVEGAVYLPGLDCLNTPCNLGCTCYPTETSYV-----CTCAPGYSGDQCEL 3089  
 QY 134 DADSIINSIDKOGAKLLGP 154  
 DB 3090 DFDCHSNPCRNAGATCVDGFN 3110

RESULT 13  
 CDNL\_HUMAN STANDARD; PRT; 164 AA.  
 AC P38936;  
 DT 01-FEB-1995 (rel. 31, Created)  
 DT 01-FEB-1995 (rel. 31, Last sequence update)  
 DT 15-DEC-1998 (rel. 37, Last annotation update)  
 DE CYCLIN-DEPENDENT KINASE INHIBITOR 1 (MELANOMA DIFFERENTIATION  
 DE ASSOCIATED PROTEIN 6) (MDA-6) (P21) (CDK-INTERACTING PROTEIN 1).  
 GN CDKN1A OR CDKN1 OR CIP1 OR WAF1 OR MDK6 OR SDI1 OR P1C1 OR CAP20.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE-94061996; PubMed-8242751;  
 RA Harper J.W., Adam G.R., Wei N., Keyomarsi K., Ellledge S.J.;  
 RT "The p21 Cdk-interacting protein Cipl1 is a potent inhibitor of G1  
 RT cyclin-dependent kinases.";  
 RL Cell 75:805-816(1993).  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-94061997; PubMed-8242752;  
 RA El-Deiry W.S., Tokino T., Velculescu V.E., Levy D.B., Parsons R.,  
 RA Trent J.M., Lin D., Mercer W.E., Kinzler K.W., Vogelstein B.;  
 RT "WAF1, a potential mediator of p53 tumor suppression.";  
 RL Cell 75:817-825(1993).  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-94081955; PubMed-8259214;  
 RA Xiong Y., Hannon G.J., Zhang H., Casso D., Kobayashi R., Beach D.;  
 RT "p21 is a universal inhibitor of cyclin kinases.";  
 RL Nature 366:701-704(1993).  
 RP SEQUENCE FROM N.A.  
 RA Jiang H., Fisher P.B.;  
 RT "Use of a sensitive and efficient subtraction hybridization protocol  
 RT for the identification of genes differentially regulated during the  
 RT induction of differentiation in human melanoma cells.";  
 RL Mol. Cell. Differ. 1:285-299(1993).  
 RP SEQUENCE FROM N.A.  
 RA Jiang H., Lin J., Herlyn M., Kerbel R.S., Weissman B.E.,  
 RA Welch D.R., Fisher P.B.;  
 RT Submitted (MAY-1994) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE-94170884; PubMed-8125163;  
 RX Noda A., Ning Y., Venable S.F., Pereira-Smith O.M., Smith J.R.;  
 RT "Cloning of senescent cell-derived inhibitors of DNA synthesis using  
 RT an expression screen.";  
 RL Exp. Cell Res. 211:90-98(1994).  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-95384154; PubMed-7655464;  
 RA Mouses S., Oezcelik H., Lee P.D., Malkin D., Bull S.B.,  
 RA Andrusis I.L.;  
 RT "Two variants of the CIP1/WAF1 gene occur together and are associated  
 RT with human cancer.";  
 RL Hum. Mol. Genet. 4:1089-1092(1995).  
 RP SEQUENCE FROM N.A.  
 RA Palmer S.;  
 RT Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.  
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 139-160.  
 RX MEDLINE-97015085; PubMed-8861913;  
 RA Gulbis J.M., Kelman Z., Hurwitz J., O'Donnell M., Kuriyan J.;  
 RT "Structure of the C-terminal region of p21(WAF1/CIP1) complexed with  
 RT human PCNA.";  
 RL Cell 87:297-306(1996).  
 CC -1- FUNCTION: MAY BE THE IMPORTANT INTERMEDIATE BY WHICH P53 MEDIATES  
 CC ITS ROLE AS AN INHIBITOR OF CELLULAR PROLIFERATION IN RESPONSE TO  
 CC DNA DAMAGE. MAY BIND TO AND INHIBIT CYCLIN-DEPENDENT KINASE  
 CC ACTIVITY, PREVENTING PHOSPHORYLATION OF CRITICAL CYCLIN-DEPENDENT  
 CC KINASE SUBSTRATES AND BLOCKING CELL CYCLE PROGRESSION.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- TISSUE SPECIFICITY: IS EXPRESSED IN ALL ADULT HUMAN TISSUES,  
 CC WITH 5-FOLD LOWER LEVELS OBSERVED IN THE BRAIN.  
 CC -1- INDICATION: BY THE P53 TUMOR SUPPRESSOR/ONCOGENE. ALSO BY MEZEREIN  
 CC (ANTIHELMINTIC COMPOUND) AND BY IEN-BETA.  
 CC -1- SIMILARITY: THE N-TERMINAL OF CIP1 AND KIP ARE SIMILAR.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: L25610; AAA16109.1; -  
DR EMBL: S67388; AAB29246.1; -  
DR EMBL: U09579; AAB55641.1; -  
DR EMBL: U03106; AAC04313.1; -  
DR EMBL: L26165; AAA19811.1; -  
DR EMBL: L47233; AAB59560.1; ALT\_INIT.  
DR EMBL: Z85996; CAB06656.1; -  
DR PIR: S39357; S39357.  
DR SWISS-2DPAGE; P38936; HUMAN.  
DR MIM: 116899; -  
KM Cell cycle; Nuclear protein; zinc-finger.  
FT ZN-FING 13 41 C4-TYPE (POTENTIAL).  
FT DOMAIN 141 156 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
SQ SEQUENCE 164 AA; 18119 MW; 98D1E7C519ADFC9 CRC64;

Query Match 8.3%; Score 71.5; DB 1; Length 164;  
Best Local Similarity 25.2%; Pred. No. 1.7;  
Matches 39; Conservative 10; Mismatches 53; Indels 53; Gaps 6;

QY 30 PAG-----GRPADRRFRPTNLSVSPWAYRISYDPARYPRYLPEAYCLCRGC----- 77  
DB 4 PAGDVQRNPGCSKACRRFLFGPVDSQLS-----RDCDALMACIQEAR 46  
QY 78 -----LTGLFGEEDVFRSAP-VYMPYVLRTTPACAGRSVYT----- 115  
DB 47 ERWNDFEYTERPLESGDFEMERNVRLGLPLKLYLPTGPRGRDELGGRRPGTSPALLQCTA 106  
QY 116 -EAYVTIPGCTCVPPEPKAD-SINSSIDKOGAK 148  
DB 107 EEDHVDLSLCTLVPRSGEQAGSGPGDSDGSRK 141

RESULT 14  
YMA0\_MYCTU STANDARD; PRT; 265 AA.  
ID YMA0\_MYCTU  
AC Q10522;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE HYPOTHETICAL 27.5 KDA PROTEIN RV2240C.  
GN RV2240C OR MTCY427.21C.  
OS Mycobacterium tuberculosis  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
CC Actinomycetales; Corynebacteriaceae; Mycobacterium.  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-H37RV;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Davlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,  
RA Honysky T., Jagers K., Kirogh A., McLean J., Moule S., Murphy L.,  
RA Rulver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Sulston J.E.,  
RA Taylor K., Whitehead S., Batteil B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence."  
RL Nature 393:537-544(1998).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL: Z70692; CA94661.1; -  
DR TUBERCULIST; RV2240C; -  
KM Hypothetical protein; Transmembrane.  
FT TRANSMEM 89 109 POTENTIAL.  
SQ SEQUENCE 265 AA; 27529 MW; F9E7ACFB3736B90A CRC64;

Query Match 8.2%; Score 70.5; DB 1; Length 265;  
Best Local Similarity 25.4%; Pred. No. 3.8;  
Matches 44; Conservative 15; Mismatches 69; Indels 45; Gaps 8;

QY 28 SCPAGRP---ADRRFRPTNRSV-----SPWAYRISYD-----PARYPR 65  
DB 24 ACCLDGRRPIVPHRRRRRIALRSVLRMRDPRPARSRCDQVTSYAVLIGRAVPRRHGG 83  
QY 66 YLPEAYCLCRGC-----LTGLFGEEDVY-----FRSAPVMPYVLRTTPACAGG----- 110  
DB 84 ELPRGALALGCIALLLNGIVGCTTVDGTAMPDTNVAAYRASSVSASVSAATSSIRE 143  
QY 111 ---RSVYTEAYVTIPGCTCVPPEPKDA-DSINSSIDKOGAKLLIGPDAPA 158  
DB 144 SQRQSLTTTKA---IRTSCDALAATSKDAIDKVNVAFAFNOGRNTGPTDEGA 193

RESULT 15  
GLYC\_NEUCR STANDARD; PRT; 479 AA.  
ID GLYC\_NEUCR  
AC P34898;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 01-FEB-1994 (Rel. 28, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE SERINE HYDROXYMETHYLTRANSFERASE, CYTOSOLIC (EC 2.1.2.1) (SERINE  
DE METHYLASE) (GLYCINE HYDROXYMETHYLTRANSFERASE) (SHMT).  
GN FOR.  
OS Neurospora crassa.  
OC Eukaryota; Fungi; Ascomycota; Sordariales; Sordariaceae; Neurospora.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92195285; PubMed=1532227;  
RA McClint C.R., Davis C.R., Denome S.A.;  
RT "Characterization of the formate (for) locus, which encodes the  
RT cytosolic serine hydroxymethyltransferase of Neurospora crassa."  
RL Mol. Cell. Biol. 12:1412-1421(1992).  
CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.  
CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETHETRAHYDROFOLATE + GLYCINE +  
CC H(2)O = TETRAHYDROFOLATE + L-SERINE  
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY)  
CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,  
CC HORMONES AND OTHER COMPONENTS.  
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO FORMS OF THE ENZYME: A  
CC CYTOSOLIC ONE AND A MITOCHONDRIAL ONE.  
CC -1- SIMILARITY: BELONGS TO THE SHMT FAMILY.  
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL: M81918; CAB26403.1; -  
DR PIR: A42241;  
DR INTERPRO: IPR001085; -  
DR PFAM: PF00464; SHMT; 1.  
DR PROSITE: PS00096; SHMT; 1.  
KM Transferrase; Pyridoxal phosphate; One-carbon metabolism.  
FT BINDING 249 249 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
SQ SEQUENCE 479 AA; 52821 MW; 06E196B035DE1E4 CRC64;

Query Match 8.2%; Score 70.5; DB 1; Length 479;  
 Best Local Similarity 19.7%; Pred. No. 7.5;  
 Matches 38; Conservative 23; Mismatches 59; Indels 73; Gaps 7;

QY	9	LSAFHHTLQLPREQARNASC-----PAGGRPADDRRRFP	42
DB	93	LEAFH----LDPKQMGVNYQCLGSPANLQVYQAIMPVHGRMLGLDLPHGHLSHGYQTP	148
QY	43	PTNLRSVSPW----AYRISYDPARYPRYLPEAYCLCRGCLTGLFGEEDYFRSAPYMP	98
DB	149	QRKISAVSYTFETMPYRVNID-----TGLI-DYDTLEKNAQLFRPK	188
QY	99	VYLRRTPA-----CAGRSYTEAYVTIPVGCVCPEPEKADINSID	143
DB	189	VLVAGTSAICRLIDYERMRKIADSVGAYLVDMAHISGLIASEVTPSPFLYADVVTTH	248
QY	144	KOGAKLLGPND	156
DB	249	KS----LRGPRGA	257

Search completed: February 16, 2001, 12:53:01  
 Job time: 191 sec

THIS PAGE BLANK (USPTO)







```

DB      522 GVEPATDADALLESHOE 540

RESULT  9
O9XUH9  PRELIMINARY;      PRT;      226 AA.

ID      09XUH9
AC      01-NOV-1999 (TREMblrel. 12, Created)
DT      01-NOV-1999 (TREMblrel. 12, Last sequence update)
DE      01-MAY-2000 (TREMblrel. 13, Last annotation update)
GN      ZK39.8 PROTEIN.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC      Rhabditidae; Pelodierinae; Gaenorhabditis.
OX      NCBI_TaxID=6239;

RP      SEQUENCE FROM N.A.
RX      MEDLINE=94150718; PubMed=7906398;
RA      Kershaw J.;
RT      "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL      elegans."
RN      Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RP      SEQUENCE FROM N.A.
RX      MEDLINE=94150718; PubMed=7906398;
RA      Bonfield J., Alnough R., Anderson K., Baynes C., Berks M.,
RA      Craxton M., Dear S., Du Z., Durbin R., Favell A., Fulton L.,
RA      Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA      Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA      Lightning J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA      Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA      Smaildon N., Smith A., Sonnenhammer E., Staden R., Sulston J.,
RA      Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA      Watson A., Weinstock L., Wilkinson-Sproat J., Woldman P.;
RT      "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RL      elegans."
RT      Nature 368:32-38(1994).
DR      EMBL; 282093; CAB05021.1;
DR      INTERPRO: IPR001304;
DR      PFM: PF00059; Jctln_c1
DR      PROSITE: PS50041; C-type lectin_2; 1.
SQ      SEQUENCE 226 AA; 24150 MW; DB4C40BFF904200 CRC64;

Query Match          9.7%; Score 83.5; DB 5; Length 226;
Best Local Similarity 25.6%; Pred. No. 0.47;
Matches 32; Conservative 18; Mismatches 34; Indels 41; Gaps 8;

OY      26 NASCPAGRPADRRPRPTNLRSVSPWARYRISYDAPRRYRLPAPYCLGRG-LTGIFGE 84
DB      52 NRGCPAGW-----TRFNRS-----GCMCVRF--PGTYHOFHLSRCQSGAVLTGVQNO 100
OY      85 EDVRSAPVYMPYV-----LRRTPACAGRSVYTEAVYTPVGTCTVPEPEKDA 135
DB      101 EEAR-KIASLLLPISQSGSIYIGLHRTAPACSKS-----PISSC----- 140
OY      136 DSINS 140
DB      141 NSMNS 145

RESULT  10
O9VFD4  PRELIMINARY;      PRT;      805 AA.
AC      09VFD4;
DT      01-MAY-2000 (T-EMblrel. 13, Created)
DT      01-MAY-2000 (T-EMblrel. 13, Last sequence update)
DT      01-JUN-2000 (T-EMblrel. 14, Last annotation update)
DE      CG18442 PROTEIN.
GN      CG18442.
OS      Drosophila melanogaster (Fruit fly).

OC      Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC      Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC      Ephydroidea; Drosophilidae; Drosophila.
OX      NCBI_TaxID=7227;

RP      SEQUENCE FROM N.A.
RX      STRAIN=BERKELEY;
RX      MEDLINE=20196006; PubMed=10731132;
RA      Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA      Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA      George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA      Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.,
RA      Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chumpe M., Pfeiffer B.D.,
RA      Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA      Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA      Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA      Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA      Borliva D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA      Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA      Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA      de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA      Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA      Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA      Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA      Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA      Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA      Jastil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,
RA      Klamel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA      Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA      Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA      Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,
RA      Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclad J.M.,
RA      Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese H.G.,
RA      Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA      Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA      Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA      Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA      Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA      Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA      Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA      Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA      Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT      "The genome sequence of Drosophila melanogaster."
RT      Science 287:2185-2195(2000).
DR      EMBL; AE003707; AAF5126.1;
DR      FLVBASE; FBgn0038287; CG18442.
SQ      SEQUENCE 805 AA; 86321 MW; FDF9FEFF14E69957 CRC64;

Query Match          9.3%; Score 80.5; DB 5; Length 805;
Best Local Similarity 21.5%; Pred. No. 3.8;
Matches 42; Conservative 24; Mismatches 60; Indels 69; Gaps 10;

OY      1 NSARAR-----AVLAFNHTLQGP-----RQANACSPGGR---PADRRPRPT 44
DB      434 NSVQAKREFQSVANGNGNITVSHSHITETATVERKSRFSCTPSKREQPGOEILPPP 493
OY      45 NLRVSPWARYRISYDAPRP-----RYLPEAYCLRGCTGLGEGEDYRFPSPAPYV 96
DB      494 R---TPTEQLSPRPAPRPKSAELLQRYSP-----KKQVRIASPYMCP 534
OY      97 -----PTVLRRTTPACAGRSVYTEA-----VTIVGCTCVPEPEKADSIINSI 142
DB      535 QERRELCPPQLPPGSGPLLDGSSQSSPTNAVSGKPLPLPIACR--PRPSNGVNSPNS-- 591
OY      143 DKQAKALLGPNAP 157
DB      592 -----SPGSAP 597

RESULT  11

```



09UBG7  
ID 09UBG7 PRELIMINARY: PRT: 520 AA.  
AC 09UBG7;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)  
DE TRANSCRIPTION FACTOR RBP-L.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PANCREAS;  
RA Koyama K., Isaka S., Okamura S.;  
RT "12-pan.";  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Isaka S., Koyama K., Nakamura Y., Okamura S., Azuma C., Kimura T.;  
RT "Human RBP-L.";  
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB024964; BAA88232.1; -  
DR EMBL: AB026048; BAA86121.1; -  
SQ SEQUENCE 520 AA; 57038 MW; BACF92459A77ACFB CRC64;

Query Match 9.2%; Score 79.5; DB 4; Length 520;  
Best Local Similarity 28.3%; Pred. No. 3;  
Matches 28; Conservative 12; Mismatches 36; Indels 23; Gaps 4;  
QY 54 YRISYDPARYPRYLPEAYCLRCGLTGLGEEYVRSASAVYPTVLRTPACAGS--- 110  
DB 426 YRGGVPAVLYCVVPDVAFC-----SDMRRLRAPITIPMSLVR-----ADGLFY 470  
QY 111 ---RSVYEAAYVTIPYGTCTVPEPEKADSIINSIDKQ 145  
DB 471 PSAFSFTYTPETYSVRP-GHGVPEPATDADALLESTHQE 508

RESULT 12  
046545 PRELIMINARY: PRT: 1045 AA.  
AC 046545;  
DT 01-JUN-1998 (TREMBlrel. 06, Created)  
DT 01-JUN-1998 (TREMBlrel. 06, last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, last annotation update)  
DE COMPLEMENT RECEPTOR TYPE 2.  
GN CR2.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=WHITE ALPINE;  
RA Hein W.R., Dudler L., Marston W., Landsverk T., Young A., Avila D.;  
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF038131; AAB92375.1; -  
DR HSSP: P10998; 1YVD  
DR INTERPRO: IPR000436; -  
DR PFM: PF00084; sushi; 14.  
SQ SEQUENCE 1045 AA; 115561 MW; PF58E1A2892C0D59 CRC64;

Query Match 9.0%; Score 77.5; DB 6; Length 1045;  
Best Local Similarity 25.9%; Pred. No. 10;  
Matches 29; Conservative 13; Mismatches 43; Indels 27; Gaps 5;  
QY 42 PPTNLRSVPMAYRISYDPARYPRYLPEAYCLRCGLTGLGEEYVRSASAVYPTVLRTPACAGS--- 99  
DB 423 PPKLNGQKEDRRHVRDPDGTISRY-----SCDPGYLVGEESIRCTPPDGVWLPFTA 473

QY 100 VLRRTPACAGRSVTEAYVTIPVGTCTVPEPEKD--ADSIINSIDKQAKL 149  
DB 474 PICKAAEC-----PVGKQVKKPKKNGPIRPDVNSCCD-EGYRL 511  
RESULT 13  
09XW14 PRELIMINARY: PRT: 354 AA.  
AC 09XW14;  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, last sequence update)  
DE Y51HA.7 PROTEIN.  
GN Y51HA.7.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94150718; PubMed=7906398;  
RA Smye R.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.";  
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94150718; PubMed=7906398;  
RA Smye R.;  
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.";  
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.

QY 20 PREGARNASCPAGRPADRRFPPTNLRSVSPMAYRISYDPARY-PRYLPEAYCLRCGL 78  
DB 141 PERRPRRSAPAS--PAKPYRPPQPARSHBEYQRIKMKVKEKMPRRRDAYDLOO--- 195  
QY 79 TGLFGEEDVFRSAPVMPYVLRTPACAGRSVTEAYVTIPVGTCTVPEPEKADSI 138  
DB 196 ---GGOOSRIQ---MLRQTLRLRAP-----LEEERK---PPVATMPVPDE----- 231  
QY 139 NSSI--DKQAGKLLGPNDAACP 160  
DB 232 HKALPWTBPGKRRKMGGEAATAP 255  
RESULT 14  
09VTX9 PRELIMINARY: PRT: 733 AA.  
AC 09VTX9;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, last annotation update)  
DE CG10663 PROTEIN.

Query Match 8.9%; Score 77; DB 5; Length 354;  
Best Local Similarity 25.0%; Pred. No. 3.6; Length 354;  
Matches 36; Conservative 20; Mismatches 56; Indels 32; Gaps 8;  
QY 20 PREGARNASCPAGRPADRRFPPTNLRSVSPMAYRISYDPARY-PRYLPEAYCLRCGL 78  
DB 141 PERRPRRSAPAS--PAKPYRPPQPARSHBEYQRIKMKVKEKMPRRRDAYDLOO--- 195  
QY 79 TGLFGEEDVFRSAPVMPYVLRTPACAGRSVTEAYVTIPVGTCTVPEPEKADSI 138  
DB 196 ---GGOOSRIQ---MLRQTLRLRAP-----LEEERK---PPVATMPVPDE----- 231  
QY 139 NSSI--DKQAGKLLGPNDAACP 160  
DB 232 HKALPWTBPGKRRKMGGEAATAP 255  
RESULT 14  
09VTX9 PRELIMINARY: PRT: 733 AA.  
AC 09VTX9;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, last sequence update)  
DT 01-OCT-2000 (TREMBlrel. 15, last annotation update)  
DE CG10663 PROTEIN.

GN CG10663.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Suton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champs M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Barker I.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Bertman B.P., Bhandari D., Bolshakov S.,  
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B., Dunn P.,  
RA Durbin K.J., Evangelista C.D., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fodor C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glöckner A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Mei M.-H., Ibbegan C.,  
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kenton J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lal Z.,  
RA Lako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy E., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinett K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson R.C., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhou G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.W., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster";  
RL Science 287:2185-2195(2000).  
DR EMBL: AE003541; AAF4916.1; -  
DR HSP: P00750; IRTF.  
DR FLIASE; FBgn0036287; CG10663.  
DR INTERPRO: IPR000884; -  
DR INTERPRO: IPR001254; -  
DR INTERPRO: IPR001314; -  
DR INTERPRO: IPR002465; -  
DR INTERPRO: IPR002965; -  
DR PRAM: PF00089; trypsin; 1.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR PRINTS: PR01217; PRICHETENSIN.  
DR PROSITE: PS00134; TRYPsin\_HIS; 1.  
DR PROSITE: PS00135; TRYPsin\_SER; 1.  
DR PROSITE: PS00340; RECEPTOR\_CYTOKINES-2; UNKNOWN\_1.  
SQ SEQUENCE 733 AA; 84355 MW; BE47F6A30E8CB61B CRC64;

Query Match 8.9%; Score 76.5; DB 5; Length 733;  
Best Local Similarity 26.1%; Pred. No. 9.1;  
Matches 36; Conservative 15; Mismatches 42; Indels 45; Gaps 8;

QY 26 NASCPAGRP-ADRR---FRPPTILRSVSPWARYISYD---PARYRPLPAAYCICRCCL 78  
DB 134 NAAAAAASGPVVDYRYEYKRPYASISYPTHTPIYTPRAHPRPQPTTY----- 186

QY 79 TGLFGEEDVFRSAPVYMT-----VLRRTAPACAGRSYTYEAYTVIPGCTCV 128  
DB 187 -----PPRTHTPTPTPTPTNDEYVRLPKYP-GKDLFSEEDVDEL----- 228  
QY 129 PEPEDADINSIDKOG 146  
DB 229 -EPEDEDLIS--DEOG 243  
RESULT 15  
Q9PL24 PRELIMINARY; PRT; 242 AA.  
AC Q9PL24;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, last sequence update)  
DT 01-OCT-2000 (TREMBLrel. 15, last annotation update)  
DE HYPOTHEITICAL PROTEIN TC0284.  
GN TC0284.  
OS Chlamydia muridarum.  
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.  
OX NCBI\_TaxID=83560;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MOB / NIG;  
RX MEDLINE=20150255; PubMed=10684935;  
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,  
RA White O., Hickey E.K., Peterson J., Uitterback T., Berry K.,  
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C.,  
RA Gwinn M., Nelson C.W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,  
RA Eisen J., Fraser C.M.;  
RT "Genome sequences of Chlamydia trachomatis Mobn and Chlamydia pneumoniae AR39.";  
RL Nucleic Acids Res. 28:1397-1406(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Read T.D., Brunham R., Shen C., Gill S.R., Heidelberg J.F., White O.,  
RA Hickey E.K., Peterson J., Umayam L.A., Uitterback T., Berry K.,  
RA Baas S., Linher K., Weidman J., Khouri H., Craven B., Bowman C.,  
RA Dodson R., Gwinn M., Nelson W., Deboy R., Kolonay J., McClarty G.,  
RA Salzberg S.L., Eisen J., Fraser C.M.;  
RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AE002296; AAF39152.1; -  
DR TIGR: TC0284; -  
KW Hypothetical protein.  
SQ SEQUENCE 242 AA; 26914 MW; 0621776E89B3BC84 CRC64;

Query Match 8.8%; Score 76; DB 2; Length 242;  
Best Local Similarity 26.2%; Pred. No. 3;  
Matches 21; Conservative 6; Mismatches 21; Indels 32; Gaps 3;

QY 68 PEAYCICRGCLTGLF-----GEEDVFRSAPVYMTVVLKRTAPACAGRSYTT--- 115  
DB 161 PEPHCNCLHCQIGRAVEEDIEVSEEDLTFERSWDI-----SQGKMYVTVD 208  
QY 116 -----EAYVIPGCTC 127  
DB 209 PLNDEGEENVVLGPIGCTC 228

Search completed: February 16, 2001, 12:52:42  
Job time: 197 sec



THIS PAGE BLANK (USPTO)

```

QY 1361 aagtggaacgtgacatcttgcagctgtcagaagaatccaagcaggtattgcttact 1420
DB 269 AAGTGAACGTGACATCTTTTCCAGTTTCTCAGAAAGATCCAGACAGGATTGGCTTACTT 210
QY 1421 gtaagggcttaagatcagctgaatgagagaagaagtgagcagcttaagctgag 1480
DB 209 GTAAGGGCTTAAAGTACAGCTGATATAGAGACAAAGTGGCCACGTTGACATCTGCAG 150
QY 1481 agatcaatcggagagctctgtcttcctcattctgcacagagagcttagcttactt 1540
DB 149 AGATCAATCTGAGAGCTTCTGTCTCTGATCTCTCCACAGAGAGCTAGGCTCTGATCTT 90
QY 1541 tcttaagtgaaagctgtctctgcacacacattattgttaaaagtgtagtcttctt 1600
DB 89 TCTTTAATGAAAGTGTCTCTCTGAAACACAAATTTGTAAGTAGTCTCTTTT 30
QY 1601 taaatcaataaagagctgtcgtga 1626
DB 29 TAAATCATTAAGAGGCTTGTCTGA 4

RESULT 4
HDMCH13C/c HDMCH13C 296 bp mRNA PRI 24-MAR-1994
DEFINITION Homo sapiens (clone ch13CDNA193) mRNA sequence.
ACCESSION L30108
VERSION L30108.1 GI:463092
KEYWORDS
SOURCE Homo sapiens female 3 month post-natal brain cdna to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 296)
AUTHORS Bonaldo,M., Soares,M.-B. and Warburton,D.
TITLE Selection of chromosome-specific cDNAs and their corresponding
genomic clones
JOURNAL Unpublished (1993)
FEATURES
source location/Qualifiers
1..296
/organism="Homo sapiens"
/db_xref="taxon:9606"
/dev_stage="3 month post-natal"
/sex="female"
/tissue_type="brain"
/map="13q11"
misc_feature 1..296
BASE COUNT 84 a 66 c 73 g 73 t
ORIGIN
Query Match 17.9%; Score 294.4; DB 85; Length 296;
Best Local Similarity 99.7%; Pred. No. 4e-36;
Matches 295; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1001 aacatgcttttactctctcgtgtagaatttttaagaacataattggaatcccttgatga 1060
DB 296 ACATGCTTTTACTCTCTCTGTGTAAGTAATTTGTAACATAATGGAATGGAATCTTGATTA 237
QY 1061 atttgtagctgtacacatcctgcgctgtgtctgaattcagctgtacacagtgctga 1120
DB 236 ATTGTTGAGCTGTACACTCTGCGCTGTCTGTGAATTCAGCTGTACCGATGCGCTGA 177
QY 1121 ctgtagtaagagacagcttcatctgacccacatctccttcacatgaagtgctcacagg 1180
DB 176 CTGATGAAATGAGACACGCTCATCTGACCCACCTCTCCTTCACATGAAGTCTTCACAGG 117
QY 1181 cctcagtgtagcaacaagagtgacaggggctgcagatgccacagggccagctaaagt 1240
DB 116 CCTCAGAGTGACCAAAAGGATGACAGGCGCTGCGATGCCACAGGGCCAGCTAAGAGT 57
QY 1241 tccaaagatctcagattggttttagcatgaatacaataaagctccaactcgc 1296

```

```

DB 56 TCCAAAGATCTCAGATTGTTTGTATGATGATATACATAAAGCTCTCAAACTGCC 1
RESULT 5
LOCUS HUMNOTIA 323 bp mRNA PRI 13-DEC-1993
DEFINITION Human chromosome specific mRNA.
ACCESSION L23206
VERSION L23206.1 GI:434047
KEYWORDS
SOURCE Homo sapiens female 3 month post natal entire brain cdna to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 323)
AUTHORS Bonaldo,M., Soares,M.-B. and Warburton,D.
TITLE Selection of chromosome-specific cDNAs and their corresponding
genomic clones
JOURNAL Unpublished (1993)
COMMENT Chromosome 13q11.
FEATURES
source location/Qualifiers
1..323
/organism="Homo sapiens"
/db_xref="taxon:9606"
/dev_stage="3 month post natal"
/sex="female"
/tissue_type="entire brain"
BASE COUNT 41 a 120 c 119 g 43 t
ORIGIN
Query Match 15.4%; Score 252.8; DB 85; Length 323;
Best Local Similarity 99.2%; Pred. No. 9.9e-30;
Matches 254; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 21 gccgtgctcagtgcttccacacacagctgagctggggcgagagagcgagcgaac 80
DB 256 GCGGTGCTCAGTGCCTTCCACACACAGCTGACGTGGGGCGGTGAGCAGGCGCAAC 197
QY 81 ggcagctccggcagggagggcagggccgcagcgcgcgttcgcgcgcacacacg 140
DB 196 GCGACCTGCCCGGAGGGGGCAGGCCCGCGGACCGCGGCTTCGGACCCACCAACTG 137
QY 141 cgcagcgtgtgcgcttgccgtgacagatctctcagacccggcgaggtacccagctac 200
DB 136 CCGACGCTGTGCGCTGTGGGCTTACAGAAATCTCTACGACCCGCGAGGTACCCAGGTAC 77
QY 201 ctgcctgaagcctactgctgtgcggggcgctgacccggcgagctgttcggcgagagagac 260
DB 76 CTGCCTGAAGCTACTGCTGTGCGGGGCTGCGGCGGCTGCTGCGGAGAGAGAC 17
QY 261 gtgcgcttcgcagcgt 276
DB 16 GTGCGCTTCCGACAGG 1

RESULT 6
LOCUS G31083 250 bp DNA PRI 31-DEC-1996
DEFINITION human STS SHGC-19175, sequence tagged site.
ACCESSION G31083
VERSION G31083.1 GI:1758484
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 250)
AUTHORS Myers,R.M.
JOURNAL Unpublished (1996)
COMMENT
Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)

```

THIS PAGE BLANK (uspto)